Talaromyces atroroseus

Genome sequencing, Monascus pigments and azaphilone gene cluster evolution

Rasmussen, Kasper Bøwig

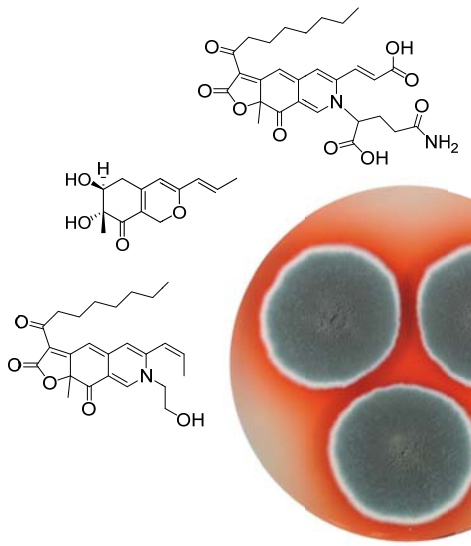
Publication date:
2015

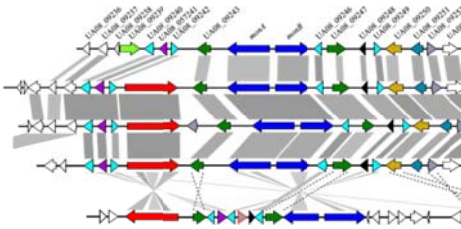
Document Version
Publisher's PDF, also known as Version of record

Technical University of Denmark

TALAROMYCES ATROROSEUS

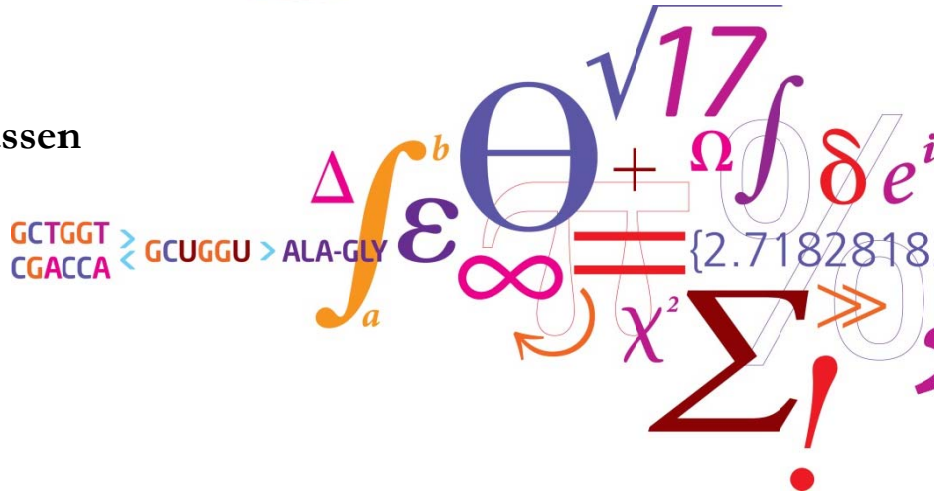
GENOME SEQUENCING, *MONASCUS* PIGMENTS AND AZAPHILONE GENE CLUSTER EVOLUTION





Kasper Bøwig Rasmussen

PhD Thesis
September 2015



Talaromyces atroroseus

Genome sequencing, *Monascus* pigments and azaphilone gene cluster evolution

By

Kasper Bøwig Rasmussen

A thesis submitted in partial fulfillment of the degree of
Doctor of Philosophy

at

Department of Systems Biology
Technical University of Denmark

September 2015

Supervisors:
Associate Professor Ulf Thrane
Professor Uffe Hasbro Mortensen

Summary

Monascus pigments have been used for food colouring in Asia for centuries. However, industrially produced *Monascus* pigments are not approved for use in foods in neither Europe nor the United States. This is partly due to potential risk of co-production of the mycotoxin citrinin by the *Monascus* species.

Alternative species for production of *Monascus* pigments have been described. These are *Talaromyces* species previously belonging to the *Penicillium* subgenus *Biverticillium* but recently transferred to *Talaromyces*. In the present PhD thesis *Talaromyces atroroseus* is examined as a suitable *Monascus* pigment producing organism. As part of the examination, one strain was selected for genome sequencing to unravel the genetic basis for *Monascus* pigment production in *T. atroroseus*.

Talaromyces atroroseus is described as a new species, separated from other *Monascus* pigment producing *Talaromyces* species. The taxonomic description is based on multigene phylogenies of ITS, β-tubulin and *RPB1*. Furthermore *T. atroroseus* is distinguished from *Talaromyces purpurogenus* by the lack of production of any known mycotoxins such as the rubratoxins and rugulovasins, which are produced by *T. purpurogenus*. The lack of mycotoxin production makes *T. atroroseus* a particular interesting species for a potential industrial pigment production.

Talaromyces atroroseus IBT11181 was chosen as model-organism for metabolic profiling of produced pigments. It was shown to produce a broad range of red *Monascus* pigments in reactor-based controlled fermentations. One of these pigments has been structure elucidated and shown to be the glutamine derivative of a known *Talaromyces* produced *Monascus* pigment, PP-V.

Talaromyces atroroseus IBT11181 was full genome sequenced thereby providing the first *Talaromyces* genome of a suitable *Monascus* pigment producer. When locating the *Monascus* pigment gene cluster, it was found that the polyketide synthase (PKS) was missing. Knock-out of the mitorubrin PKS identified this as the PKS delivering the azaphilone backbone to both the group of *Monascus* pigments as well as the mitorubrins. This form of genetic entanglement of biosynthesis of two groups of secondary metabolites leading to loss of the PKS from one gene cluster is to my knowledge a new observation in the fungal kingdom. This finding indicates that the machinery behind the vast diversity of fungal natural products is much more complex than the genetic machinery, within the normal gene cluster boundaries.

By analysis of the *Monascus* pigment and the mitorubrin gene cluster in *Talaromyces* I identified five homologous gene pairs to be present in both clusters. Phylogenetic studies of the evolutionary relationship between these gene pairs from the *Monascus* pigment and the mitorubrin gene cluster in *Talaromyces* and *Monascus* give strong evidence that the two clusters originate from a deep segmental duplication of an ancestral azaphilone gene cluster in a common ancestor to *Talaromyces*, *Monascus* and *Aspergillus* lineages. The mitorubrin cluster has since been lost in *Monascus*, while it has evolved to the azanigerone cluster in *Aspergillus niger*. In *T. atroroseus,* the azaphilone PKS from the *Monascus* pigment gene cluster has since been lost. The loss is probably a result of redundancy in product formation by the two azaphilone PKSs from the *Monascus* pigment and mitorubrin gene clusters. Utilizing the mitorubrin PKS in production of *Monascus* pigments might have given the fungus a selective advantage leading to fixation of the *Monascus* pigment PKS loss.

Further studies of the mitorubrin gene cluster led to identification of highly homologous gene clusters in genome sequenced *Stachybotrys* species. Gene-by-gene phylogenies groups several of the *Stachybotrys* cluster genes adjacent to the *Talaromyces* mitorubrin cluster genes and within other *Eurotiomycetes* genes, although the genus *Stachybotrys* belongs to distant related *Sordariomycetes*. This phylogenetic study hence gives strong hints towards a horizontal gene transfer of the mitorubrin cluster from a *Talaromyces* ancestor to a *Stachybotrys* ancestor.

It is very valuable that the gene clusters for azaphilones and other pigments are widely found within the fungal kingdom. This gives numerous possibilities to unravel and understand the detailed genetic regulation of pigment production and thereby establishing a sound platform for safe fungal cell factories producing colorants.

Dansk Resume

Monascus pigmenter har i Asien været brugt til farvning af fødevarer i århundreder. Industrielt producerede *Monascus* pigmenter er dog imidlertid ikke godkendt til brug i fødevarer i hverken EU eller USA. Dette skyldes til dels risikoen for at samproduktion af mykotoksinet citrinin når *Monascus* arter anvendes som produktionsorganismer.

Alternative arter til produktion af *Monascus* pigmenter er tidligere blevet identificeret. Disse alternative produktionsorganismer tilhører den tidligere *Penicillium* underslægt *Biverticillium,* men er for nylig blevet taksonomisk omklassificeret til slægten *Talaromyces*. I denne Ph.d. afhandling bliver arten *Talaromyces atroroseus* karakteriseret og undersøgt for dens egnethed om *Monascus* pigment produktionsorganisme. Som led i dette er én *T. atrorosues* stamme blevet udvalgt til fuldgenom sekventering for at udforske det genetiske grundlag for *Monascus* pigment produktion i *T. atroroseus*.

Talaromyces atroroseus beskrives som en ny art. Den adskiller sig fra andre *Monascus* pigment producerende *Talaromyces* arter på baggrund af fylogenetiske analyser af ITS, β-tubulin og *RPB1*. Derudover adskiller *T. atroroseus* sig fra *Talaromyces purpurogenus* ved ikke at producer kendte mykotoksiner så som rubratoksiner og rugulovasiner. At *T. atroroseus* ikke producer nogen kendte mykotoksiner gør arten til en attraktiv potential pigment producent set fra et industrielt perspektiv.

Talaromyces atroroseus IBT11181 blev udvalgt som model til at studere sammensætningen af producerede pigmenter. Det viste sig at stammen producerer en bred vifte af røde *Monascus* pigmenter når den dyrkes under kontrollerede forhold i bioreaktorer. Et af de producerede

pigmenter blev strukturopklaret. Det viste sig at være en variation af det kendte *Talaromyces* producerede *Monascus* pigment PP-V, men med glutamin bundet i azaphilone ringstrukturen.

Talaromyces atroroseus IBT11181 blev fuldgenom sekventeret. Dette er det første *Talaromyces* genom af en egnet og effektiv producent af *Monascus* pigmenter. Genklustret ansvarlig for produktion af *Monascus* pigmenter blev identificeret, men viste sig overraskende at mangle selve polyketid syntasen (PKS). PKS'en fra mitorubrin klustret blev slået ud og det viste sig at denne PKS er ansvarlig for at producere den bicykliske kernestruktur i både *Monascus* pigmenterne og mitorubrinerne. Denne genetiske sammenfiltring af biosyntesen for to separate grupper af sekundære metabolitter, førende til tab af PKS'en fra det ene genkluster, er efter min viden en ny observation i svamperiget. Dette fund er med til at indikere at det genetiske maskineri, som danner grundlag for den store diversitet af naturprodukter fra svampe, er meget mere komplekst og at produkterne ikke kun er forbundet med generne indenfor de normale grænser af genklustrene.

Analyse af genklustrene til produktion af *Monascus* pigmenter samt mitorubriner i *Talaromyces,* afslørede at fem gener har en homolog i det andet kluster. Fylogenetiske analyser af den evolutionære sammenhæng mellem disse homologe gen-par blev foretaget, inkluderende gener fra *Monascus* pigment genklustret i *Monascus* arter. Resultaterne viser at de to genklustrer stammer fra det samme azaphilone genkluster, som har undergået en duplikation i en fælles forfader til *Talaromyces, Monascus* og *Aspergillus*. Mitorubrin-klustret er sidenhen gået tabt i *Monascus* mens det har udviklet sig til azanigeron klustret i *Aspergillus niger*. I *T. atroroseus* er PKS'en fra *Monascus* pigment genklustret gået tabt. Dette er sandsynligvis sket som en konsekvens af redundansen i at have to store PKS gener der i sidste ende producerede identiske eller tæt på identiske azaphilone strukturer. Brugen af mitorubrin PKS'en i produktion af *Monascus* pigmenter har sandsynligvis givet svampen en selektiv fordel, og tabet af den oprindelige *Monascus* PKS er som resultat blevet fikseret i *T. atroroseus*.

Under yderligere analyser af mitorubrin genklustret i *T. atroroseus* blev det opdaget at fuldgenom sekventerede *Stachybotrys* arter har et genkluster med høj homologi til mitorubrin-klustret i *Talaromyces* arterne. Fylogenetiske analyser af mitorubrin-kluster generne grupperede flere af generne fra *Stachybotrys* arterne som søster-grupper til mitorubrin generne fra *Talaromyces* arterne. Disse *Stachybotrys* gener falder fylogenetisk indenfor gener fra andre *Eurotiomycetes,* selvom *Stachybotrys* tilhører fjernt beslægtede *Sordariomycetes*. De fylogenetiske træer for metabolit generne er således ikke i overensstemmelser med arts-fylogenien og de

præsenterede analyser giver derfor stærke antydninger om at mitorubrin-klustret har været horisontalt overført fra en *Talaromyces* forfader til en *Stachybotrys* forfader.

Udbredelsen af azaphilone genklustre på tværs i svampe riget er særdeles værdifuldt. Det giver mulighed for en dybere forståelse for den genetiske regulering af produktionen af disse pigmenter og derigennem danne grundlaget for etablering af sikre cellefabrikker til produktion af farvestoffer.

Preface

With this thesis I conclude my time as PhD student at DTU Systems Biology. The work presented was conducted between August 2011 and June 2015 and was supervised by Associate Professor Ulf Thrane and Professor Uffe Hasbro Mortensen with financial support by DTU and grant 09-064967 from the Danish Council for Independent Research, Technology, and Production Sciences.

The last four years has been great, but at times also tough and demanding with lots of challenges and ups and downs. One of the biggest challenges has probably been to cover over many technical and scientific areas introducing me to the world of analytical chemistry as well as bioinformatics and at the same time building upon my previous experience in molecular microbiology. Within the field of bioinformatics I have often felt a bit alone and especially the steep learning curve within genomics, handling large data sets including next generation sequencing data, has been challenging but at the same time rewarding when things at the end finally worked out. One thing is for sure, if it were not for some of the people around me I would not have reached the point of finishing this PhD thesis.

Firstly I would like to thank my supervisor Ulf Thrane for giving me the opportunity to enhancing my scientific knowledge by conducting the PhD study at DTU Systems Biology. Ulf, you have always been available when I needed to discuss small scientific questions, project strategies or more administrative or personal matters. I will like to thank you for that. In some part of the project I have been exploring scientific areas far from the original project frame, but you have always been supportive when the project where reaching dead-ends.

I would also like to thank my co-supervisor Uffe Hasbro Mortensen for valuable discussions and guiding me in the right directions at some critical points during the project. I

want to thank you for getting me to focus the thesis at a time point where the entire project was getting a bit chaotic for me.

Furthermore I would like to thank Kristian Fog Nielsen for introducing me to the world of analytical chemistry. Your door has always been open for quick questions and problem solving, for which I am grateful.

Throughout my PhD time I have had the wonderful joy to share office with a lot of different people whom I would like to thank for delightful conversations whether it has been scientific, work-related or completely off-topic. You know who you are.

A good working atmosphere is important for the motivation to go to work. I would like to thank all my colleagues at DTU Systems Biology in building 221 and 223; none mentioned, none forgotten. You have all been part of making the workdays fun and full of valuable discussions, ideas etc.

At last I would like to thank my family. A great thank goes to my kids for bringing joy in my life when the PhD project at times deprived me from energy and motivation. Special thanks go to my lovely wife Minka Hickman. You have throughout the PhD time and especially in the last phase been patient and most of the time understanding when frustrations and work pressure in periods affected my presence at home.

Kasper Bøwig Rasmussen

September 2015

Table of Contents

List of Figures

List of Tables

Abbreviations

ACP	Acyl carrier protein
AT	Acyl transferase
bn	Billion
BPC	Base peak chromatogram
CAZy	Carbohydrate active enzymes
CDD	Conserved domain database
cMT	C-methyltransferase
CY	Czapek yeast extract
CYA	Czapek yeast autolysate agar
DAD	Diode array detector
DLTRS	Duplication, loss, transfer, rates and sequence
DTU	Technical University of Denmark
EIC	Extracted ion chromatogram
ER	Enoyl reductase
EWC	Extracted wavelength chromatogram
FAS	Fatty acid synthase
GH	Glycoside hydrolase
GTR	General time reversible
HGT	Horizontal gene transfer
HPLC	High performance liquid chromatography
ILS	Incomplete lineage sorting
kb	Kilobases
KR	Keto reductase
KS	β-keto-synthase
LBA	Long-branch attraction

LECA	Lightweight expanded clay aggregates
MCC	Maximum clade credibility
MCMC	Markov chain Monte Carlo
MP	*Monascus* pigment
MPR	Most parsimonious reconciliation
MS	Mass spectrometry
NMR	Nuclear magnetic resonance
nr	Non-redundant protein database at NCBI
NR-PKS	Non-reducing polyketide synthase
NRPS	Non-ribosomal peptide synthase
ORF	Open reading frame
PCR	Polymerase chain reaction
PE	Paired end
PK	Polyketide
PKS	Polyketide synthase
PT	Product template
RT	Retention time
SAT	Starter unit ACP transacylase
SM	Secondary metabolite
TE	Thioesterase
UHPLC	Ultra-high performance liquid chromatography
USER	Uracil specific excision reagent
UV/VIS	Ultraviolet/visible

Chapter 1
Introduction

Over the last decades consumer demands for safer and more natural products have made its mark on the nature and extent of chemicals used in plastics, textiles, foods and any other product that come in contact with humans. A subgroup of these affected chemicals is colorants and especially food colorants, also referred to as food pigments. Food colorants is a billion dollar industry valued at around $1.54 billion in 2011 and has been increasing rapidly over the last decade ($1.07bn in 2004, $1.15bn in 2007, (Mapari et al., 2010; Leatherhead, 2013)); an increase that is expected to continue according to market analysis (MarketsandMarkets, 2013). The rapid growing market value is partly explained by the ongoing paradigm shift from using synthetic colorants towards using natural or natural derived colorants; the latter two going from a market share at 34% in 2007 to 39% in 2011 (Leatherhead, 2013) and expected to steadily increase in coming years.

With the rapid increase of the market for natural food colorants it is interesting from an industrial point of view to investigate new possibilities for novel food colorants and the production of these. Several of the approved natural colorants in the European Union and the United States are extracted from natural sources such as beetroot, grapes or bell peppers (Mapari et al., 2005). However drawbacks such as pH-sensitive colouration, instability caused by prolonged exposure to light and undesirable odours and tastes make it interesting to explore new sources for natural food colorants.

The production methods of plant-derived natural pigments also have drawbacks. Extractions of these pigments are usually carried out by direct extraction from agricultural crops. The production of pigments is hence vulnerable to variation in quality and to the

supply of raw materials. These raw materials can have large seasonal variations in quantity and quality due to external factors such as weather influencing the harvest. The demand for pigments and consistency in the quality and colour hue of pigments used for food production, has driven the development towards production of pigments in filamentous fungi (Dufossé et al., 2014).

Monascus pigments (MPs) produced by filamentous fungi belonging to the *Monascus* genus have been used for food colouring in Asia for centuries but are not authorized for use in foods in neither the European Union nor the United States. This is partly because of the risk for co-production of the nephrotoxic mycotoxin citrinin (Liu et al., 2005; Dufossé et al., 2014). However potential mycotoxin-free cell factories for the production of MPs have previously been identified (Mapari et al., 2009). These alternative MP producers belong to *Penicillium* subgenus *Biverticillium,* which has recently been transferred to the *Talaromyces* genus (Samson et al., 2011).

This increasing interest for improving production methods for natural food colorants as well as the identification of alternative *Monascus* pigment producing strains forms the basis of the present PhD thesis.

1.1 Aim of the study

My PhD project is part of ongoing research of pigment production by filamentous fungi at Section of Eukaryotic Biotechnology at DTU Systems Biology. More specifically my focus has during the last years been specifically on MP production by *Talaromyces atroroseus*. The work on pigment production has been conducted as part of collaborative effort with PhD student Gerit Nymschefsky to investigate and further develop MP production using *T. atroroseus* as production organism. It builds upon previous work in screening and identifying potential pigment producing fungal cell-factories performed by PhD Sameer Mapari (Mapari, 2008) at the former Center for Microbial Biotechnology (now Section for Eukaryotic Biotechnology) at Department of Systems Biology, Technical University of Denmark.

The aim of my part of the project can be divided in two. First part was to support the fermentative production of *Monascus* pigments with analytical tools for chemical fingerprinting and potential quantification of produced pigments. The other, and major part, was to genome sequence *T. atroroseus* to unravel the biosynthetic basis for azaphilone pigment production, thereby creating a basis for improving *Monascus* pigment production by genetic engineering.

1.2 Structure of the thesis

Throughout the thesis figures and tables are numbered by the chapter number followed by the number within the corresponding chapter, e.g. Figure 5.4 is figure number 4 from Chapter 5. Appendices are assigned with letters and appendix figures and tables are represented by the appendix letter followed by the table of figure number, e.g. Figure D.12 is figure 12 from Appendix D.

The thesis is structured in eight chapters where the present **Chapter 1** introduces the background for the project as well as aims and structure of the thesis.

Chapter 2 provides some historical and theoretical background to support the following chapters. This includes a brief introduction to pigments. Furthermore filamentous fungi and their secondary metabolites are introduced with focus on biosynthesis of polyketides as well as a general introduction to azaphilones and *Monascus* pigments. Furthermore methods for phylogenetic inference are briefly introduced.

Chapter 3 comprises the published manuscript "*Talaromyces atroroseus*, a new species efficiently producing industrially relevant pigments", which is the taxonomic description of *Talaromyces atroroseus,* separating the species from other red pigment producing *Talaromyces.*

In **Chapter 4** initial steps towards describing the pigment producing abilities of *Talaromyces atroroseus* IBT11181 is taken. This is done in a controlled reactor-based sub-merged fermentation setup. My aim of the study was investigate the MP profile of *T. atroroseus* by submerged fermentation and hence no effort in order to evaluate and optimise production conditions has been taken by me. This will however be covered by Gerit Nymschefsky's PhD project. In this chapter it is shown that a broad range of *Monascus* pigment was produced and most of these of unknown structure according to the applied dereplication procedure. One of the compounds was purified and structure elucidated to be a novel variation of the already known *Monascus* pigments.

In **Chapter 5** I present the genome sequence of *Talaromyces atroroseus* IBT11181. The genome sequence is used to identify the MP gene cluster, which showed to have lost the most important gene encoding the polyketide synthase. Knockout of the mitorubrin PKS gene however resulted in abolishment of MP production, which suggests an intertwined biosynthesis of the two groups of azaphilones in *T. atroroseus.*

The finding in **Chapter 5** leads to questions addressing the putative evolutionary link between the two azaphilone clusters. In **Chapter 6** this evolutionary relationship and potential

origin of the azaphilone gene clusters within *Eurotiales* is unravelled by phylogenetic inference of gene tree reconciliations.

In **Chapter 7** a gene cluster in *Stachybotrys* is found to have high homology to the mitorubrin cluster from *Talaromyces* species. Phylogenetic analysis suggests that an ancestral gene cluster was horizontal transferred between *Talaromyces* and *Stachybotrys*.

Finally in **Chapter 8** the conclusions from the PhD thesis is summarised and future perspectives are discussed.

1.3 References

Dufossé, L., Fouillaud, M., Caro, Y., Mapari, S.A. & Sutthiwong, N., **2014**. Filamentous fungi are large-scale producers of pigments and colorants for the food industry. *Current Opinion in Biotechnology*, **26**: 56–61. doi:10.1016/j.copbio.2013.09.007

Leatherhead, **2013**. *New research reveals natural colours overtake artificial / synthetic colours for first time*. Press release, 28 February, Leatherhead Food Research, Surrey, viewed 25 June 2015, <https://www.leatherheadfood.com/new-research-reveals-natural-colours-overtake-artificial-synthetic-colours-for-first-time>

Liu, B.-H.H., Wu, T.-S.S., Su, M.-C.C., Chung, C.P. & Yu, F.-Y.Y., **2005**. Evaluation of citrinin occurrence and cytotoxicity in *Monascus* fermentation products. *Journal of Agricultural and Food Chemistry*, **53**: 170–175. doi:10.1021/jf048878n

Mapari, S.A.S., **2008**. *Chemotaxonomic exploration of fungal biodiversity for polyketide natural food colorants*. PhD thesis, Technical University of Denmark.

Mapari, S.A.S., Meyer, A.S., Thrane, U. & Frisvad, J.C., **2009**. Identification of potentially safe promising fungal cell factories for the production of polyketide natural food colorants using chemotaxonomic rationale. *Microbial Cell Factories*, **8**: 24. doi:10.1186/1475-2859-8-24

Mapari, S.A.S., Nielsen, K.F., Larsen, T.O., Frisvad, J.C., Meyer, A.S., et al., **2005**. Exploring fungal biodiversity for the production of water-soluble pigments as potential natural food colorants. *Current Opinion in Biotechnology*, **16**: 231–238. doi:10.1016/j.copbio.2005.03.004

Mapari, S.A.S., Thrane, U. & Meyer, A.S., **2010**. Fungal polyketide azaphilone pigments as future natural food colorants? *Trends in Biotechnology*, **28**: 300–307. doi:10.1016/j.tibtech.2010.03.004

MarketsandMarkets, **2013**. *Natural and synthetic food color market worth $2.3 billion by 2019*.Press release, MarketsandMarkets, Vancouver, USA, viewed 25 June 2015, <http://www.marketsandmarkets.com/PressReleases/food-colors.asp>

Samson, R.A., Yilmaz, N., Houbraken, J., Spierenburg, H., Seifert, K.A., et al., **2011**. Phylogeny and nomenclature of the genus *Talaromyces* and taxa accommodated in *Penicillium* subgenus *Biverticillium*. *Studies in Mycology*, **70**: 159–183. doi:10.3114/sim.2011.70.04

Chapter 2
Background

2.1 Pigments

Colorant, pigment or dye? These terms are often used interchangeably for substances responsible for colouration of its matrix. Strictly speaking, however, a distinction between pigments and dyes, which are both covered by the term colorant, is made as pigments are insoluble in its matrix and dyes are soluble in its matrix. Hence a colorant can change from being a dye to being a pigment if the matrix is changed from being e.g. hydrophobic to hydrophilic. For biological pigments however this distinction is normally not used, and hence throughout this thesis the term pigment will be used to cover all colouring compounds.

Pigments are compounds capable of absorbing visual light thereby changing the colour of the reflected light. The reflected light is then observed by the human eye as the colour of the light not absorbed. Hence a pigment absorbing light in the low wavelengths of visual light (~380-530 nm) will appear red and a pigment with absorption in the high wavelengths of visual light (~600-700 nm) will appear blue and so forth.

The ability for pigments to absorb light, and thereby appear coloured, is linked to the chemical structures and more specifically to the systems of conjugated double bonds. When double bonds appear in a conjugated system the electrons can delocalise across the system of overlapping p-orbitals. When presented to light the electrons can absorb photons of specific wavelengths depending of the extent and nature of the conjugated system. This results in only some wavelengths being reflected resulting in observation of coloured compound by the human eye. In Figure 2.1 the structures of some characteristic natural pigments are shown

clearly visualising the systems of conjugated double bonds. In general larger conjugated systems absorb photons with higher wavelength leading to different colour hues. Conjugated systems can both be linear conjugated systems as observed in the orange β-carotene, or aromatic or cyclic conjugated systems, as seen in anthraquinones such as carminic acid (see Figure 2.1). The colouration by some pigments like the chlorophylls also depends on chelated metal ions.

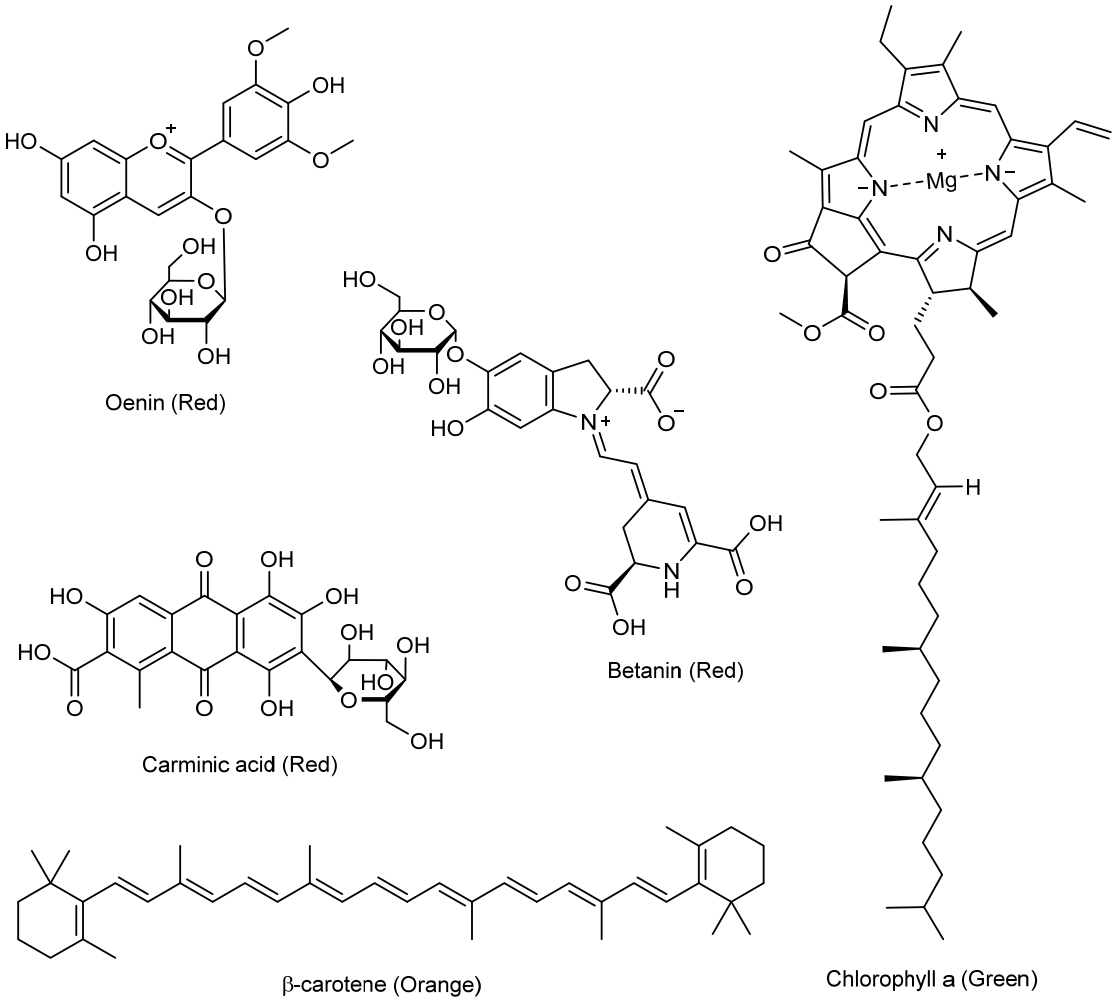


Figure 2.1. Selected characteristic natural pigments. Chlorophylls are responsible for the green colour in plants and involved in the absorbance of light in the photosynthetic process. Oenin is an anthocyanin one of the largest pigment groups in the plant kingdom. Oenin exhibits the red/purple colour in i.e. grapes. Betanin is belonging to the class of betalains responsible for yellow/orange and red/violet colouring in a variety of vegetables. Betanin is especially known for the deep colour in beetroot. β-carotene has antioxidant activity and in the liver it is broken down to a form of Vitamin A. It is responsible for the orange colour in e.g. carrots, sweet potatoes and cantaloupe melons.

2.2 Filamentous fungi

The fungal kingdom consists of eukaryotic organisms and is distinct from the kingdoms of plants and animals. The kingdom is estimated to encompass around 1.5 million species (Hawksworth & Rossman, 1997). The majority of fungi can non-scientifically be grouped in three major groups; yeasts, moulds and mushrooms. However the true taxonomical grouping is a bit more detailed. In 2007 a great effort to Assembling the Fungal Tree of Life (AFTOL, http://aftol.org/) resulted in the grouping of the Kingdom Fungi into seven phyla as well as the subkingdom *Dikarya* (Hibbett et al., 2007). The majority of described fungal species belong to the subkingdom *Dikarya,* which can be further split in the two phyla; *Basidiomycota* and *Ascomycota* (see Figure 2.2). The fungi belonging to *Dikarya* share the ability to create dikaryotic hyphae, which are hyphae with two nuclei. The *Basidiomycetes* are growing as hyphae and producing their sexual spores in *basidiocarps* often occurring as large fruit bodies known from mushrooms.

The other phylum *Ascomycota* is known as sac fungi, which is further sub-divided in three subphylums. The subphylum *Saccharomycotina* contains the true yeasts, or budding yeasts, including the historical and industrial important *Saccharomyces cerevisiae.* It also contains the genus *Candida*, which includes opportunistic human pathogens, and many more. *Taphrinomycotina* is another subphylum containing, among others, the fission yeast *Schizosaccharomyces pombe.* The third and last subphylum *Pezizomycotina* is also the most species-rich within *Ascomycotina.* It contains most of the fungi known as moulds, including important species such as industrial "work-horse" *Aspergillus niger* from the *Eurotiomycetes*, as well as several crop pathogens as the rice blast fungi *Magnaporthe grisea* and the head blight fungi *Fusarium graminearum*, both belonging to the *Sordariomycetes*. Throughout this thesis filamentous fungi will refer to the filamentous ascomycetes belonging to the subphylum *Pezizomycotina.*

Filamentous fungi are heterotrophs and rely on enzymatic machinery for degradation of organic matter for the utilisation in energy production and growth. Filamentous fungi cover different life styles being either saprophytic, pathogenic, symbiont or a combination and are often capable of living and surviving in a broad range of environments. Hence filamentous fungi occupy all major ecological niches.

Filamentous fungi are known for the production and potentially secretion of a wide range of small bioactive molecules, many of these with potential beneficial effects. Filamentous fungi are hence an excellent resource for discovering and producing novel pharmaceutical products such as antibiotics, immunosuppressants and anticancer drugs. Due to the excellent

arsenal of carbohydrate active enzymes and capability for rapid production and secretion of these, filamentous fungi are also an excellent resource for industrial production of enzymes.

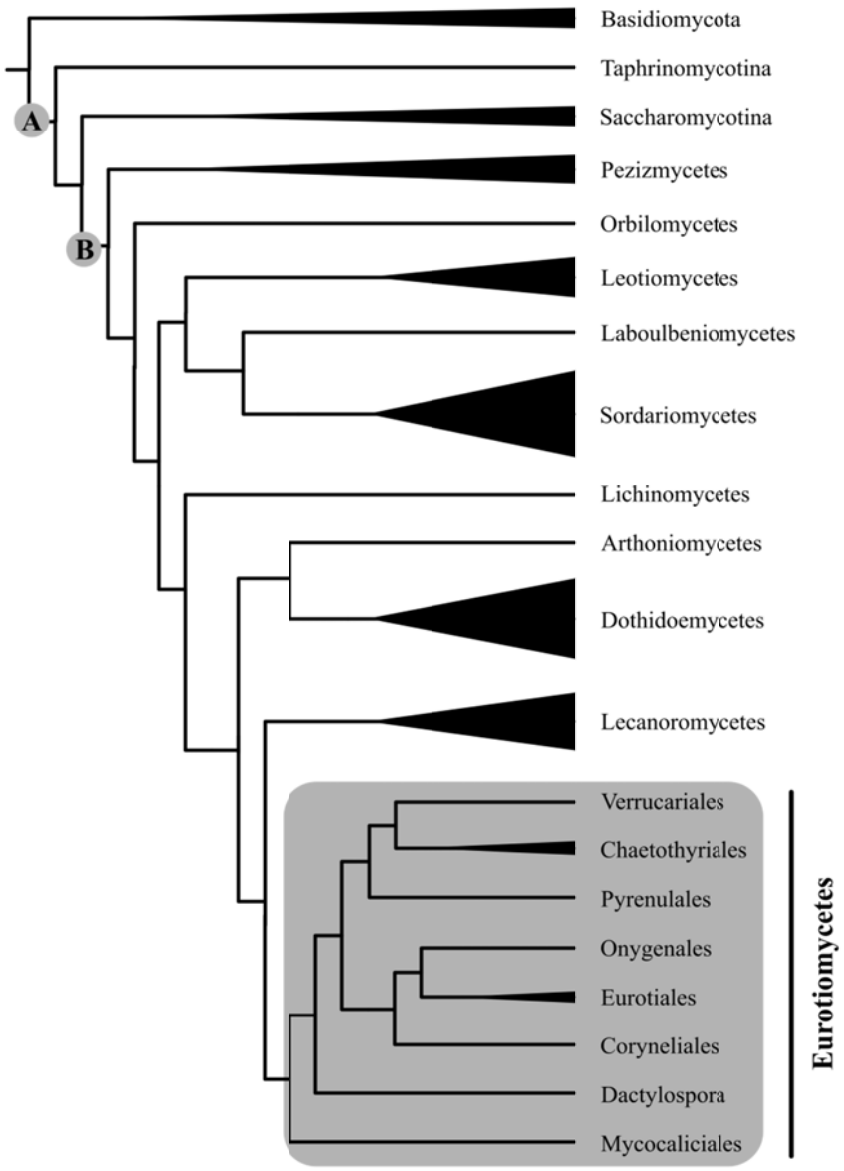


Figure 2.2. Part of the fungal tree of life covering the subkingdom *Dikarya*. The phylogenetic relationship of subphylums within *Dikarya* with more details on *Eurotiomycetes* is shown. The figure is made based on the reconstructed phylogeny from large phylogenetic studies of *Ascomycota* (Hibbett et al., 2007; Schoch et al., 2009). A: *Ascomycota*, B: *Pezizomycotina*.

2.3 Secondary metabolites

Secondary metabolites (SM) are small organic compounds giving a beneficial effect for the fungi in a competitive ecological environment. The effects of SMs are only partly uncovered and include protection against DNA damage by UV light (melanins), defence mechanism towards other organisms (antibiotics, antifungals and insecticides) and some act as signalling molecules. Their production is regulated as a response to external stimuli (Brakhage, 2013) and SMs are therefore not essential for growth in the filamentous fungi. This is in contrast to primary metabolites, which are essential for fungal growth. In general, fungi prioritise growth over SM production, since SM production is energy economical expensive for the fungi. Thus, the fungi in general start to produce the major part of SMs in late exponential phase of the fungal growth. This is at least what is observed when grown under controlled laboratory conditions.

SMs are organic molecules that even though not being essential often have very important roles for the fungi. The broad spectrum of biological activities of SMs make them an very interesting area to study for the scientific community as a very large portion of new medical drugs are either SMs or synthetic/semi-synthetic compounds created by inspiration of natural occurring SMs (Newman & Cragg, 2012). Fungal SMs are hence a great resource for discovering novel bioactivities and potential new drug candidates.

The probably best known fungal SM is penicillin, which was discovered by Alexander Fleming in 1928 (Fleming, 1929) to be produced by *Penicillium rubens* as defence against *Staphylococcus*. Since then, several other fungi-produced antibiotics has been discovered and used for treatment of bacterial or fungal infections, including cephalosporins produced by *Acremonium chrysogenum* (Brakhage, 1998) and the antifungal griseofulvin produced by *Penicillium grisefulvum* (Finkelstein et al., 1996). Other important fungi derived SMs used in medical treatments include cyclosporine A, originally discovered from *Tolypocladium inflatum* and used as immunosuppressant during implant surgery (Hoffmeister & Keller, 2007); and lovastatin, which is produced by *Aspergillus terreus* as well as a number of *Monascus* species (known as Monacolin K), and works as a cholesterol-lowering agent (Tobert, 2003).

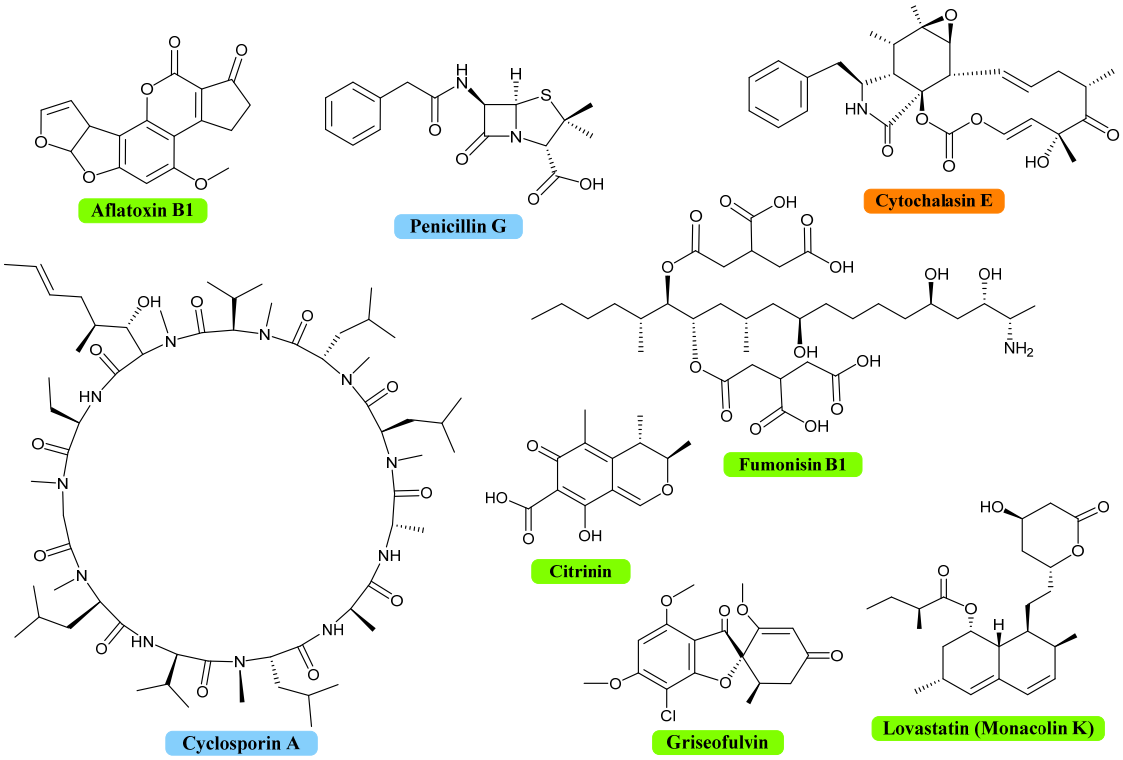


Figure 2.3. Some important secondary metabolites produced by fungi. The colouring refers to the structural nature of the compounds. Green: polyketides, blue: non-ribosomal peptides, orange: hybrids of polyketides and non-ribosomal peptides.

As fungi produce SMs to compete bacteria or other fungi they are also able to produce SMs targeting insects, such as insect juvenile hormone (Nielsen et al., 2013), or SMs with toxic effects in vertebrates. The latter are referred to as mycotoxins, which means fungal poisons. Among the major myctotoxins are aflatoxins, ochratoxins, citrinin as well as the *Fusarium* produced toxins fumonisins and trichothecenes (Yazar & Omurtag, 2008). Mycotoxins are of special concern caused by the risk of ending in prepared foods. Internationally, limits have been set for the content of some of the most occurring mycotoxins in final food and feed products.

Secondary metabolites in filamentous fungi can be divided in three major classes based on their biosynthetic nature: Polyketides (PKs), non-ribosomal peptides (NRPs) and terpenoids. Most of the studied fungal SMs belong to the PK class, which is also the compound class, which is dealt with in this thesis.

2.3.1 Polyketides and their biosynthesis

Polyketides are the largest and most abundant class of SMs in fungi. It is a group of structurally organic compounds and apart from fungal biosynthesized polyketides, bacteria,

plants and animals also produce polyketides. Even though PKs show remarkable structural diversity, the polyketide backbone for all PKs is biosynthesized in the same manner. The basic concept of PK biosynthesis involves repeated condensation of short chain carboxylic alcohols, typically in the form of acetyl coenzyme A (CoA) and malonyl CoA, to form extending carbon chains with several ketide units. Following each elongation step the resulting ketone can either stay unreduced, undergo partly reduction to create a hydroxyl group or a double bond, or be fully reduced as is the case in fatty acid synthesis.

PKs are synthetized by polyketide synthases (PKSs). These enzymes are generally divided into three classes; type I, type II and type III (Cox, 2007). Type II PKSs are usually found in bacteria and resembles the bacterial fatty acid synthase (FAS) type II. Type II PKSs consist of a system of monofunctional polypeptides each containing one functional domain of the PKS assembly line (Simpson, 1995). Type III PKSs are homodimeric enzymes usually associated with production of aromatic polyketides in plants and bacteria. Examples of type III PKS in fungi have however also been reported (Hashimoto et al., 2014). Type III PKSs, in contrast to type I and type II, do not use acyl carrier protein (ACP) domains as carriers of the extending polyketide chain. Type I PKS are large and multifunctional polypeptide enzymes, which can be further divided into the modular type I PKSs in bacteria and the iterative type I PKSs (iPKS) in fungi (Cox & Simpson, 2009).

The minimal required domains for a PKS to be functional are the ketosynthase (KS), the acetyl transferase (AT) and the acyl carrier protein (ACP) (McDaniel et al., 1994). Initially, a starter unit, often in form of acetyl CoA, is loaded to the KS domain, see Figure 2.4. The extender unit, typically in the form of malonyl CoA, is following recruited by the AT domain and loaded to the ACP domain. The C-C bond formation is hence catalysed by the KS via a decarboxylative Claisen condensation between the two thioesters (Shen, 2003), see Figure 2.4. This process is iteratively repeated until a predetermined chain length is reached, after which the PK chain is released from the PKS.

Apart from the minimal requirements of a fungal PKS additional optional domains exists adding to the chemical complexity of PKs (See Table 2.1). The β-ketoreductase (KR), enoyl reductase (ER) and dehydrogenase (DH) domains are involved in the reduction of the ketone unit during each step of elongation. If all three modules are used the ketone is fully reduced to a saturated alkyl. In the bacterial modular type I PKSs the presence of these domains in each PKS module can be used to predict the final PKS products as each domain in the module is used. However in the fungal iterative type I PKSs the catalytical activity of the domains are

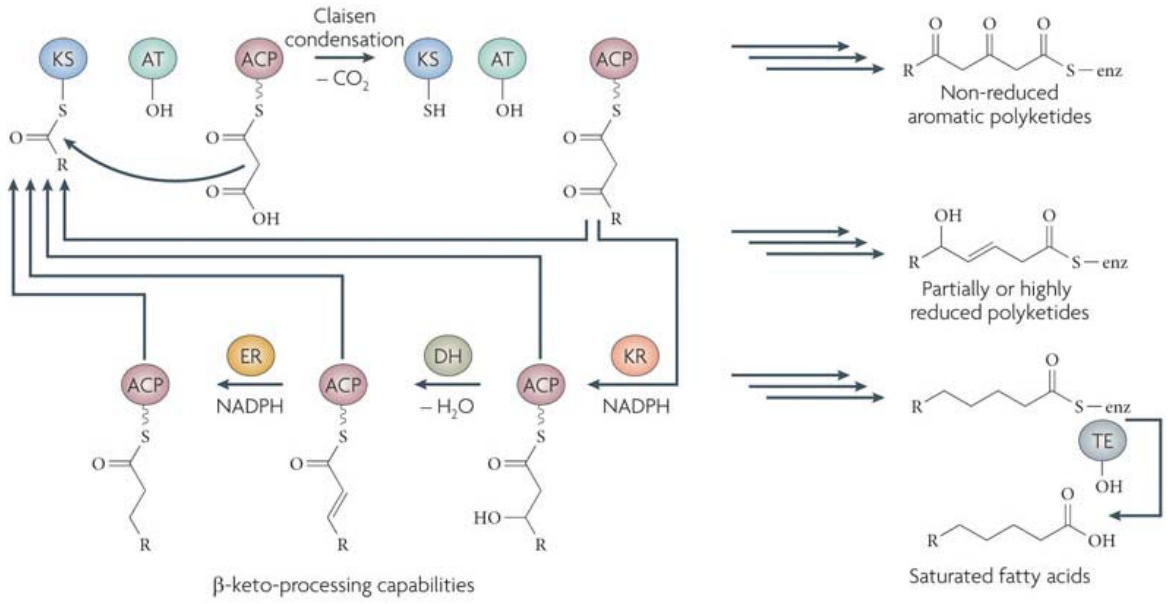


Figure 2.4. Basic principle for polyketide biosynthesis. Starter unit is loaded to the KS domain and extender unit is loaded to the ACP domain assisted by the AT domain. The KS catalyzes the condensation of the condensation reaction resulting in extrension of the PK chain. Following each extension, the ketone can be unreduced, or undergo partial or full reduction, catalysed by the KR, DH and ER domain. The figure is taken from (Crawford & Townsend, 2010).

reused for several rounds in chain elongation and each of the three reducing domains (if present in the PKS) may be used in some chain elongation cycles and not in others, thereby adding to the polyketide complexity. Understanding the programming of the fungal iterative PKSs is of great interests. Identification of structural features involved in determination of the number of iterations, or determining factors for using/not using the reducing domains in each iteration, could greatly enhance the possibilities for creation of novel PKs using synthetic biology.

Fungal PKS can be grouped in two large groups, non-reducing PKSs (NR-PKS) and reducing PKSs (R-PKS). While R-PKSs contain one or more of the three reductive domains, the NR-PKSs do not have any of the reducing domains and are involved in the production of aromatic polyketides. NR-PKSs usually contains a starter unit ACP transacylase (SAT) domain responsible for selection of alternative starter units for the polyketide biosynthesis (Crawford et al., 2006). The SAT domain is not found in R-PKSs. Furthermore fungal NR-PKSs also contain a product template (PT) domain involved in the cyclization of the polyketide chain resulting in aromatic compounds.

Table 2.1. Functional domains of polyketide synthases

Required domains	Abbreviated	Function
β-ketoacyl synthase	KS	Condensation of extender unit with growing PK chain
Acyl carrier protein	ACP	Holding the growing PK chain
Acyl transferase	AT	Loading of starter and extender acyl units
Optional domains		
β-ketoreductase	KR	Reduction of β-ketones to hydroxyls
Dehydratase	DH	Reduction of hydroxyls to enoyls
Enoyl reductase	ER	Reduction of enoyls to alkyls
C-methyl transferase	cMT	C-methylation
Starter unit ACP transacylase	SAT	Loading of starter acyl units in NR-PKS
Product template	PT	Involved in folding pattern control in NR-PKS
Claisen cyclase / Thioesterase	CLC / TE	Release of PK from PKS
Thiolester reductase	R	Reductive release domain similar to what is found in NRPSs

In addition to the domain responsible for the reduction of the ketone, some PKSs also contain a C-methyl transferase (cMT), which is responsible for addition of methyl groups during chain elongation using S-adenosylmethionine (SAM) as methyl-donor.

Fungal R-PKS usually does not contain domains specifically involved in the release of the polyketide chain but instead rely on e.g. discrete enzymes (Du & Lou, 2010). Fungal NR-PKS on the other hand often contain C-terminal release domains. These are either Claisen cyclases (CLC)/thiolesterases (TE) or thiolester reductase domains (R). The ladder is similar to release domains sometimes seen in in non-ribosomal peptide synthases (NRPS) (Du & Lou, 2010).

Fungal iterative PKSs can also be found as megasyntase hybrids with NRPS modules at the C-terminal end. The NRPS domain is then responsible for condensing amino acids to the elongated polyketide chain, thereby creating hybrid molecules.

For further in depth descriptions of PKSs and their function, the reader is referred to excellent reviews by Cox (Cox, 2007) and Hertweck (Hertweck, 2009).

2.3.2 Secondary metabolite gene clusters

The PKS itself is not responsible for the broad diversity of fungal SMs and most produced polyketides are modified upon release from the PKS creating families of the specific compound class. The post PK-synthesis modifications include, but are not limited to, O-methylations, reductions or oxidations, and attachment of other compounds to the PK structure, i.e. terpenoids, amino acids, sugars and more. The enzymes catalysing these post synthesis modifications are broadly referred to as tailoring enzymes.

This genetic machinery for biosynthesis of secondary metabolites in filamentous fungi tends to co-localize in genomic clusters (Keller & Hohn, 1997) often localized in subteleomeric regions. This is a hallmark trait of fungal SM machinery and in contrast to what is usually observed for SM biosynthetic genes in plants were the biosynthetic genes often are dispersed across the genomes (Kliebenstein & Osbourn, 2012).

In addition to genes for modification of the polyketide structure generating improved diversity several other functions are important in the biosynthetic machinery. As SMs are not essential to the host, the production of these metabolites is often linked to the response to external signals from the environment, i.e. in the response to the presence of bacteria, or as melanin production in some species is light induced, giving protection against UV light. To provide rapid response to such external factors, many SM pathways are often regulated by transcription factors (TF), which rapidly turn on necessary genes to imitate the production of SMs. Several cases of silent SM gene clusters in fungi, which were turned on by overexpression of the clustered TF, signifies the strong regulatory mechanism the fungi possess (Bergmann et al., 2007; Chiang et al., 2009; Zabala et al., 2012).

The rationale in having genes involved in production of the same class of SMs clustered in the fungal genomes has been a matter of speculation since the discovery of the first SM gene clusters. Among the most proliferated speculations is that the clustering allow for efficient transcriptional regulation due to chromatin modifications (Keller et al., 2005). Another speculation is that the physical clustering of genes facilities the possibility for rapid loss and gain of SM clusters. This can be seen as a response to ecological pressure with the gain of SM clusters facilitated by horizontal gene transfer (HGT) from other fungi living in the same ecological niche (Wisecaver & Rokas, 2015).

2.4 Azaphilones

Azaphilones are a structurally diverse group of pigmented polyketides. They are characterised by their pyranoquinone bicyclic core, which is highly oxygeneated, along with a chiral quaternary carbon at the center (See Figure 2.5). Azaphilones is known for their ability to react with amines by exchanging the pyran oxygen with nitrogen forming vinylogous γ-pyridones (Stadler et al., 1995). This exchange typically results in colour change from yellow/orange towards red. Reactions can occur both with ammonia or more complex amines such as amino acids or peptides.

The production of azaphilones is widespread in the fungal kingdom. Detailed studies on their biosynthesis has primarily been undertaken in well-established lab-organisms such as *Aspergillus nidulans* and *Aspergillus niger* (Chiang et al., 2009; Zabala et al., 2012).

The structural diversity of the azaphilones (see Figure 2.5) exhibits a broad range of biological activities (Osmanova et al., 2010). Citrinin is one of the major food-contaminating mycotoxins (Pascual-Ahuir et al., 2014) and have nephrotoxic activity towards mammals. Isochromophilones inhibit binding of the HIV envelope glycoprotein gp120 to the glycoprotein CD4 present on the surface of immune cells (Matsuzaki et al., 1995), thereby acing against HIV infections of the immune system. Comparison with inhibition by structurally close related ocrephilone and rubrorotiorin showed that few structural modifications dramatically changed the inhibition activity (Matsuzaki et al., 1995). Bulgarialactone B and related azaphilones inhibits the heat shock protein Hsp90 (Musso et al., 2010). Furthermore a range of azaphilones have antimicrobial, anticancer or anti-inflammatory activities (Osmanova et al., 2010; Gao et al., 2013).

For further structural diversity of azaphilones I refer the reader to the comprehensive review of known azaphilones published in 2013 by Gao *et al.* (Gao et al., 2013).

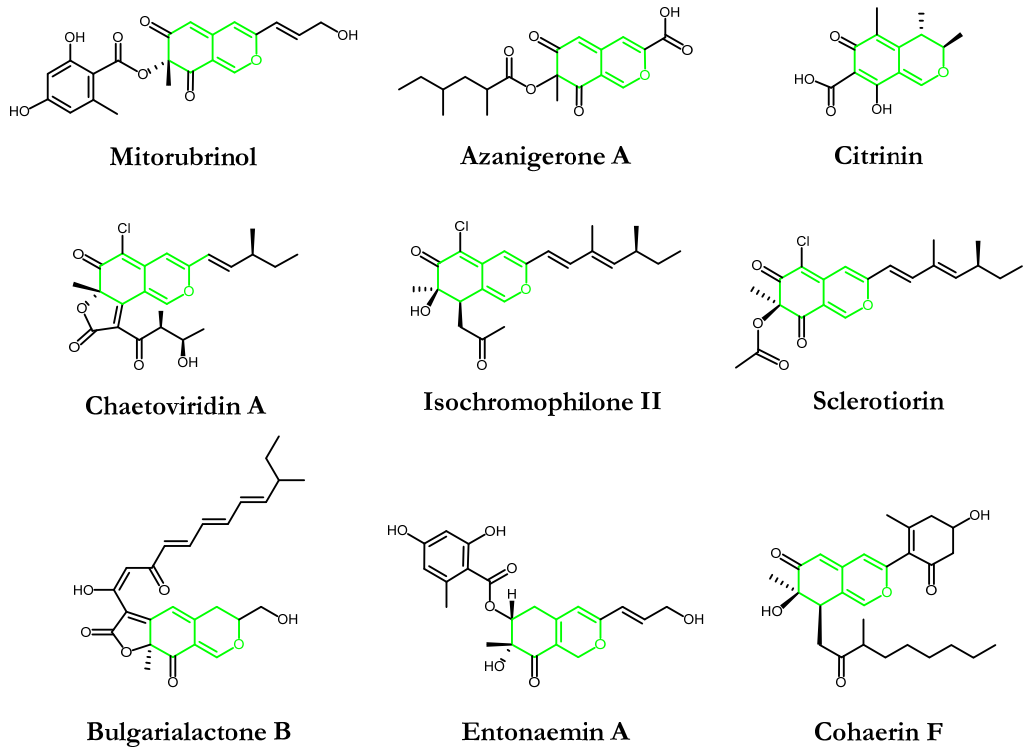


Figure 2.5. Selected fungal azaphilones. Azaphilone production is distributed widely in the fungal kingdom and the structural diversity is tremendous (Gao et al., 2013). The bicyclic core, highlighted in green, is the common structural trait of these compounds.

2.5 *Monascus* pigments

Monascus pigments (MPs) is the group of azaphilone pigments which has been most extensively studied. This is due to the historical application of MPs in food colouring in the Asian kitchen and the Asian food industry, where the tradition of using moulds for fermentative preparation of food is long. Koji is the important starting point for many traditional Asian and especially Japanese food products, i.e. sake, soy sauce, miso and rice vinegar. Koji is a mould culture prepared by growing traditionally *Aspergillus oryzae* on cooked or steamed grains or beans. The usage of Koji in the Japanese kitchen dates back to at least 300 B.C. were it was mentioned in the ritual text *Rites of Zhou* (Shurtleff & Aoyagi, 2012). One special form of koji is the red rice koji (also known as red yeast rice, Ang-khak or *hong qu*) which is made by fermentation of steamed rice with cultures of *Monascus purpureus*. The first known report of red rice koji was in it Chinese form *hong qu* in the *Chhing I Lu* (Anecdotes, Simple and Exotic) by Tao Ku and dated at A.D. 960-965 (Huang, 2000; Shurtleff & Aoyagi, 2012). One of the anecdotes describes the recipe of red pot-roasted lamb, where red rice koji is used in the simmering of the lamb meat (Huang, 2000). Since then red rice koji has been adopted and used for colouring a wide variety of Asian food products ranging from Asian wines such as the Taiwanese rice wine *Hong Ru* (Kumasaki et al., 1962) and the red variety of Japanese sake *akaisake*.

Until 1895 the fungus responsible for red pigmentation in red rice koji was known as ang-khak rice mold or simply red yeast. In 1895 however the fungi was isolated and studied by Went and given the name *Monascus purpureus* (Went, 1895). *Monascus purpureus* produces MPs as a mixture of compounds. According to references in Whalley (Whalley, 1963) the first MP, monascorubrin, was isolated in 1926 by Nishikawa. In early 1930's structural studies on the yellow MP monascin was carried out (Salomon & Karrer, 1932); however it was not before around 1960 that correct structures were assigned to some of the MPs, i.e. the orange rubropunctatin (Haws et al., 1959; Haws & Holker, 1961) and the orange monascorubrin (Fielding et al., 1960; Kumasaki et al., 1962) (see Figure 2.6). Both of these orange pigments were reported to readily react with ammonia creating the red rubropunctamine and monascorubramine respectively (Haws et al., 1959; Kumasaki et al., 1962).

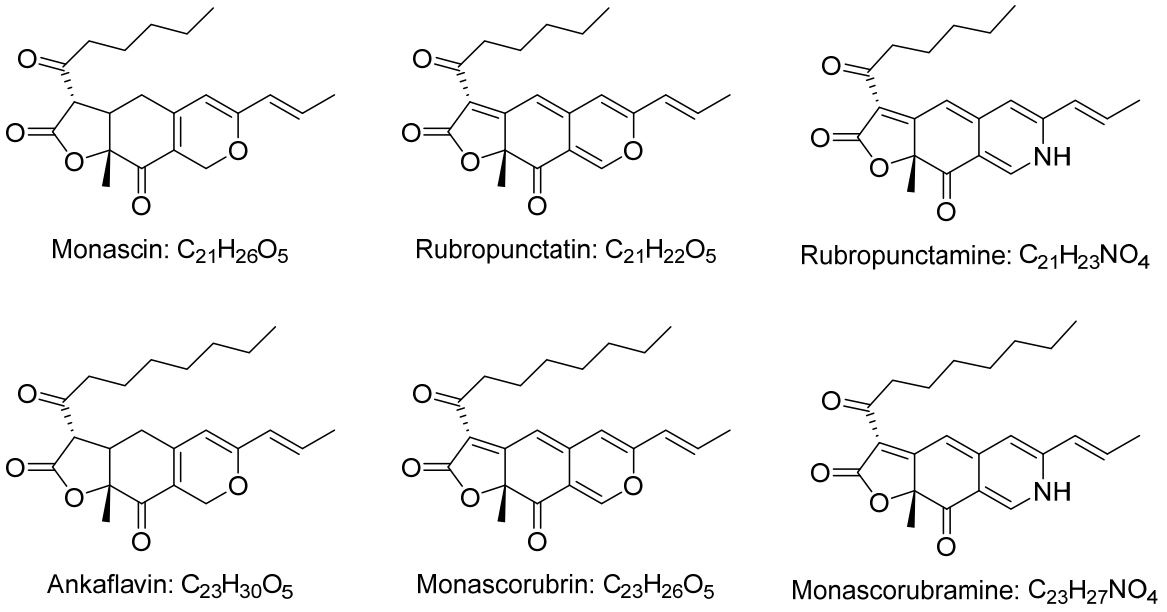


Figure 2.6. Structures of the six "original" *Monascus* pigments.

Structure of the yellow monascin was proposed in early 1960's based on a range of structural studies (Fielding et al., 1961; Whalley, 1963). At around the same time Inouye *et al.* preferring the name monascoflavin over monascin suggested an alternative structure of monascin (Inouye et al., 1962). The correct structure of monascin was in 1969 confirmed to be the one originally proposed by Fielding et al in 1961 (Chen et al., 1969).

A second yellow MP, ankaflvain, is also produced by *Monascus* species. In 1961 the structure of ankaflavin was hypothetised to be a metabolite in MP mixtures (Whalley, 1963; Fielding et al., 1961). In 1973 ankaflvain was isolated and structure elucidated by Manchand *et al.* (Manchand et al., 1973). The name ankaflavin refer to the production organism *Monascus anka* (now *Monascus purpureus*).

Collectively these six pigments are referred to as the "original" *Monascus* pigments. The six pigments have three different chromophores and hence come in pairs, see Figure 2.6. Each MP pair, having the same chromophore, also has very similar UV/VIS spectra, see Figure 2.7. The monascin and ankaflavin having the smallest chromophore are yellow compounds with λ_{max} around 390 nm, rubropunctatin and monascorubrin have λ_{max} around 470 nm, while the two red MPs monascorubramine and rubropunctamine have two characteristic smooth peaks around 415 and 525 nm.

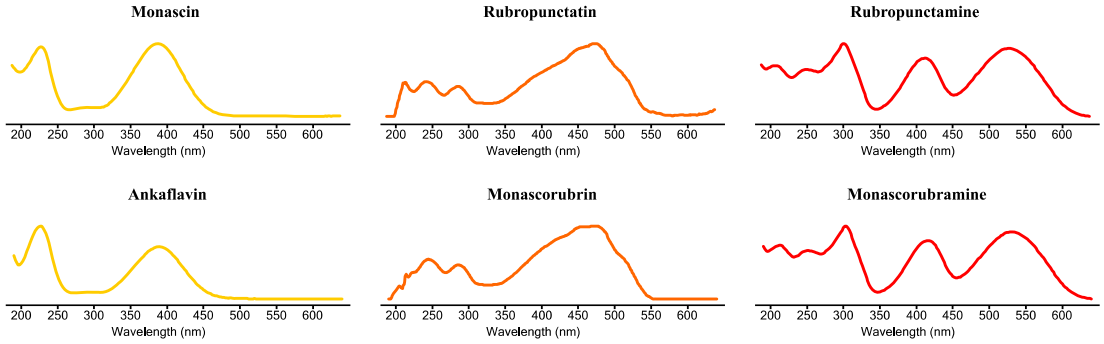


Figure 2.7. UV/VIS spectra of the six original *Monascus* pigments.

Since the structure elucidation of the six original *Monascus* pigments, many more MPs have been elucidated as produced by either *Monascus* or *Talaromyces* species. Up to June 2012 more than 50 MPs have been purified and structure elucidated (Feng et al., 2012). Since comprehensive reviews of MPs and other azaphilone pigments already exists (Osmanova et al., 2010; Feng et al., 2012; Gao et al., 2013) I will only mention some of the most significant ones in relation to this study.

In 1994 Blanc *et al.* identified two red MPs complexed with glutamate (Blanc et al., 1994) which confirms previous semi-synthetic studies (Lin et al., 1992). The two compounds have since been given the names N-glutarylmonascorubramine and N-glutarylrubropunctamine (Hajjaj et al., 1997), see the structures in Figure 2.8. Since then several studies have investigated red amino acid derivatives of MPs, their production and properties (Sato et al., 1997; Jung et al., 2003; Hajjaj et al., 2012; Lin et al., 1992). These amino acid derivative MPs are of special interest since the exhibit increased water solubility, as well as increased pH and photostability (Jung et al., 2005; Mapari et al., 2009b). These features are particular important for the potential use of novel MPs as food pigments.

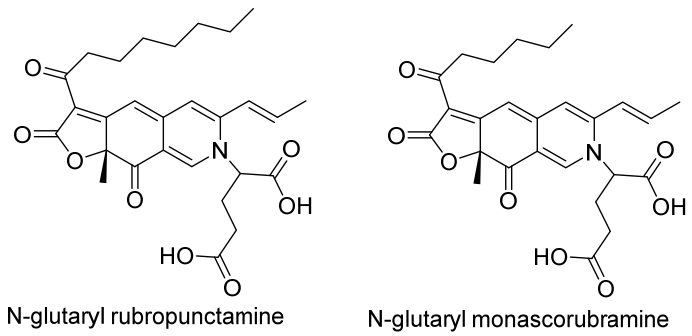


Figure 2.8. Glutamic acid derivatives of rubropunctamine and monascorubramine.

In 2004 two small azaphilones was purified from *M. kaoliang* KB20M10.2 and structure elucidated (Jongrungruangchok et al., 2004). The structure of the two azaphilones showed out to be FK17-P2b1 (wrongly named FK17-P2b2 in the publication, see more in Chapter 4) and the novel azaphilone monascusone A. Based on the two compounds a biosynthetic pathway for monascin was proposed, Figure 2.9.

The role of MPs for the fungi is yet unknown; however, the readily reaction with amines could suggest a role in nitrogen capturing when the fungi grow in a competitive environment. The MPs exhibit a range of biological activities, which also could be linked to the pigments role for the natural host. Among these activities anti-microbial (Martínková et al., 1995)(Jůzlová et al., 1996), anti-cancer (Akihisa et al., 2005; Zheng et al., 2010) and potential anti-obesity (Choe et al., 2012; Nam et al., 2014) can be mentioned. However, these activities are not necessarily desirable to have in MP mixtures for use in prepared foods. When finding novel MPs and risk assessing it is paramount to address the biological activities before considering using them as potential food colorants.

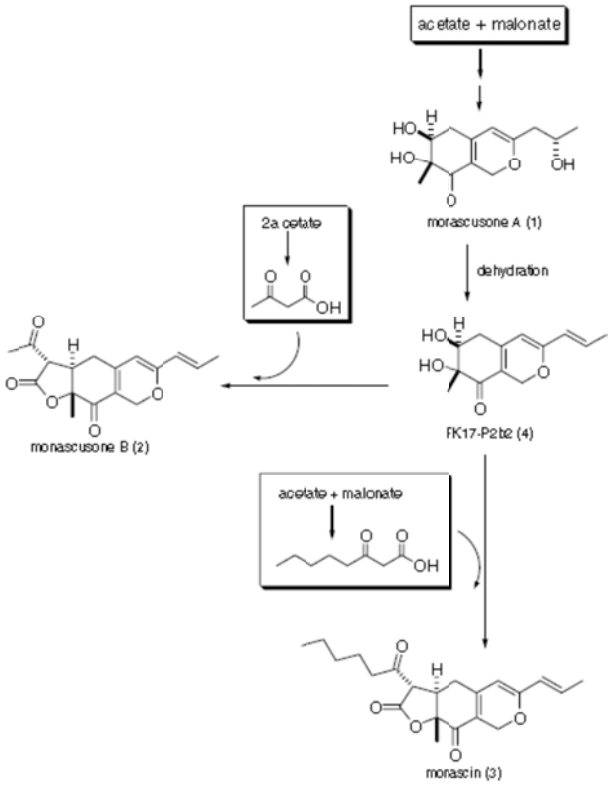


Figure 2.9. Biosynthesis of monascin as proposed by Jongrungruangchok *et al.* The biosynthesis was bassed on the detection of two small azaphilone compounds in a mutant of *Monascus kaoliang*. Figure taken from (Jongrungruangchok et al., 2004)

2.5.1 *Monascus* pigments produced by *Talaromyces* species

Since the discovery of MPs, the production of this group of compounds has also been linked to *Penicillium* species from the subgenus *Biverticillium*, which has now been transferred to the *Talaromyces* genus (Samson et al., 2011). Ogihara and co-workers reported a number of novel *Monascus* pigments produced by an unidentified *Penicillium* strain with the isolate name *Penicillium* sp. AZ (Ogihara et al., 2000a; Ogihara et al., 2000b; Ogihara et al., 2001; Ogihara & Oishi, 2002). These encompasses a total of four new MPs which were named PP-V, PP-R, PP-O and PP-Y, where PP refers to *Penicillium* pigment, and the R, V, O and Y appendix refer to the observed colour of the compounds: Red, Violet, Orange and Yellow, respectively.

PP-Y is identical to monascorubrin with the only exception that the double bond at the polyketide end is in Z-configuration instead of E-configuration as observed in monascorubrin. This difference is also a feature of the other PPs that distinguish them from known MPs produced by *Monascus* species. The stereocenters in the PPs have not been assigned with stereochemical configurations and might also be different from the MPs produced by *Monascus* species. In PP-O and PP-V the polyketide end is furthermore carboxylated, a structural feature which was not previously been observed in any MPs, but is found in mitorubrinic acid (Natsume et al., 1985), belonging to the mitorubrin class of azaphilones.

Apart from these novel MP compounds *Talaromyces* species have also been found to produce many of the MPs originally identified in *Monascus* species (Mapari et al., 2008; Mapari et al., 2009a).

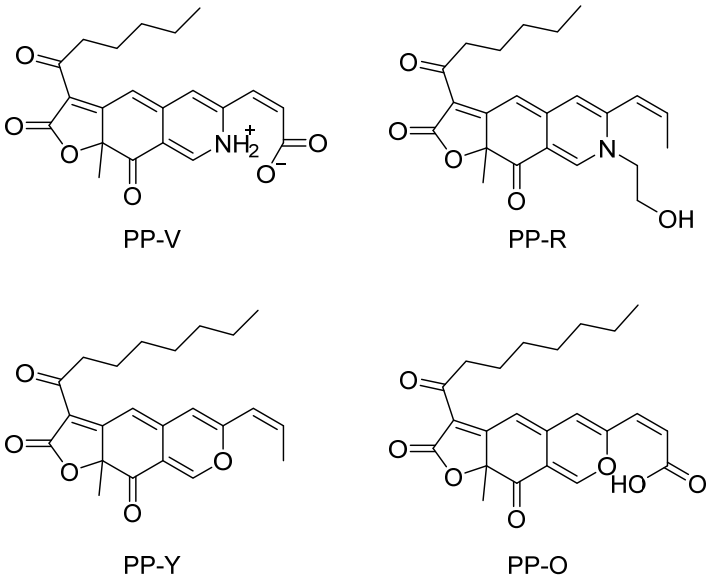


Figure 2.10. *Monascus* pigments produced by *Penicillium* sp. AZ.

2.5.2 Production of *Monascus* pigments

Monascus pigments in the have traditionally been produced by solid state fermentations (Srianta et al., 2014). More specifically cultures of *Monascus purpureus* are mixed with steamed rice, spread on large trays and preferably incubated under controlled conditions (Dufossé et al., 2005; Hsieh et al., 2008). The resulting red rice is hence sold as a moist red paste or in dried pulverized form.

However in order to develop the production methods for reaching higher yields more controlled conditions and to decrease the potential co-production of mycotoxins investigation of submerged-fermentation of *Monascus* cultures have received great attention since around 2000 (Lee et al., 2001; Domínguez-Espinosa & Webb, 2003; Gunasekaran & Poorniammal, 2008; Silveira et al., 2008). However, there is still room for optimization and together with risk of co-production of citrinin it is interesting to look towards alternative production organisms within the *Talaromyces* genus.

2.5.3 *Monascus* genomes

Studies on biosynthesis of SMs have taken a big leap since available genomes of filamentous fungi accelerated as part of the Next Generation Sequencing (NGS) revolution. However, at the initiation of the current PhD project no *Monascus* genome was public available and no linking of MP production to genes had been published. The *Monascus pilosus* genome was sequenced by the Bioresource Collection and Research Center of Taiwan in 2004, but has never been made public available (Shao et al., 2014). In 2012 the group of Chen noted in their publication, that the genome sequence of *Monascus ruber* M7 will be made available soon (Chen et al., 2012). I have nevertheless not been able to find it in public literature or public databases. The studies on biosynthesis of MP in *Monascus* published by others during my PhD time, have so far utilized these not public genomes.

As part of the 1000 Fungal Genome Project, driven by JGI with principal investigator Joseph Spatafora, two draft *Monascus* genomes were made public available in 2012 at the JGI Genome Portal. These two genomes, *Monascus purpureus* NRRL1596 and *Monascus ruber* NRRL1597, were made available prior to their publication. In 2015, the fifth (third public available) *Monascus* genome in form of the genome of *M. purpureus* YY-1 was published (Yang et al., 2015). It is available for download from the laboratory's own website (http://www2.tust.edu.cn/duxj/). Storing and maintaining genomes in small local genome

databases is in my opinion not the best way to facilitate the collectively research knowledge and progress. The research community would benefit, if genomes also are submitted to one of the large public genome databases. In this way the genomes will be easy accessible for other researchers allowing for more comprehensive studies of fungal biodiversity.

2.6 Phylogenetic inference methods

2.6.1 Phylogenetic trees

The evolutionary relationship between biological species or family of genes is usually inferred by constructing the corresponding phylogenetic tree. Data types could potentially be of many sorts ranging from morphological and phenotypic traits to molecular data in form of DNA or protein sequences. The rapid development of sequencing technologies has in recent years resulted in large amounts of molecular sequence data. This availability has resulted in molecular data has become the most widely used and accepted data-type for trustworthy inference of evolutionary relationship.

The methods and algorithms for constructing phylogenetic trees are rapidly developing. I will briefly describe the principles behind some of the most used methods.

2.6.2 Maximum parsimony

Parsimonious methods for phylogenetic tree reconstruction are the simplest to explain and "understand". Basically it relies on finding the phylogenetic tree which requires fewest events to explain the data used for the reconstruction. In simple cases with few taxa this can be done be constructing all possible trees, and then count the required events in each tree. The maximum parsimony tree is then the tree which requires least number of changes. However with increasing number of taxa the possible number of trees makes it computational intensive to construct and evaluate all trees. Therefore, most algorithms for parsimonious tree reconstruction use heuristic approaches to identify a tree that reasonably fits the data.

Maximum parsimony methods however have some drawbacks. Importantly it implies the assumption that the evolutionary model is parsimonious. One of the occurring pitfalls in phylogenetic inference using parsimony is the possible inconsistency between constructed tree and represented data that can occur due to Long Branch Attractions (LBA). LBA occurs when some taxa have long branches while other taxa have short branches. The long branches

representing a high level of substitutions can result in multiple substitutions at the same site. This gives homoplasic substitutions that in the parsimonious framework are being interpreted as a synapomorphies.

2.6.3 Maximum likelihood

Another method, which in many cases has outcompeted maximum parsimony methods is maximum likelihood (ML). As opposed to the maximum parsimony method ML requires a model of evolution. ML tries to maximize the likelihood of a hypothesis given the data. In phylogenetic context, the hypothesis will be the selected model of evolution as well as a selected tree topology. The data will often be a multiple alignment of DNA or protein seqeunces. In terms of conditional probability the likelihood can be written as:

$$Likelihood(H) = P(D|H)$$

where D is the data and H is the hypothesis. The model of evolution is used to calculate the likelihood for a given tree. In principle an exhaustive search calculating the likelihood for all possible trees could be performed. In this case the tree with the highest likelihood will naturally be found. However, as for the maximum parsimony approach, the large tree space makes these calculations a real computational effort, and hence also for ML tree construction heuristic approaches are used to compute an estimated ML tree.

2.6.4 Bayesian inference

While MP and ML are widely used in phylogenetic inference they have limitations in not always producing the phylogenetic relationships that is best decribed by the data. Bayesian inference is a probabilistic approach based on a theorem by Thomas Bayes in the 18th century. Due to computational limitations it has however not been broadly used until the rapid development of computational power in the recent decades. Bayes' theorem is given as:

$$P(H|D) = \frac{P(D|H) \cdot P(H)}{P(D)}$$

The probability of a hypothesis *H* given the data *D* is equal to the probability of the data given the hypothesis, weighted by the fraction of the probability of the hypothesis by the probability of the data. In general terms, assuming that the data have a probability of 1, Bayes'

theorem can be simplified and stated as: the posterior probability ($P(H|D)$) is proportional to the likelihood ($P(D|H)$) times the prior probability ($P(H)$). This means that prior knowledge can be incorporated into the tree construction by applying appropriate prior probability distributions of parameters. This is both the force as well as pitfall in Bayesian inference of phylogenies. If the prior probability distribution is not set realistically, the posterior probability of the hypothesis can be affected, so wrong trees will have higher probability than the right tree. One of the forces of Bayesian inference over MP and ML is that it does not give a point estimate of the phylogenetic relationship, but rather give the possibility for accounting for uncertainties in the constructed tree.

Solving the Bayesian equation is however not "doable" due to problems in computing the denominator, since this requires summation over all possible hypothesis (Felsenstein, 2004). Instead Markov chain Monte Carlo (MCMC) methods are used often in the form of the Metropolis algorithm (Metropolis et al., 1953).

The principle of MCMC methods is depicted in Figure 2.11 using the robot metaphor. Imagine a robot walking around a hilly landscape. The landscape is equal to the posterior distribution of parameters. Initially the robot is placed on a starting point. This could be either randomly selected or user defined, for example a neighbour joining tree. Using a stochastic process a new position is proposed by choosing a random direction and a gamma distributed

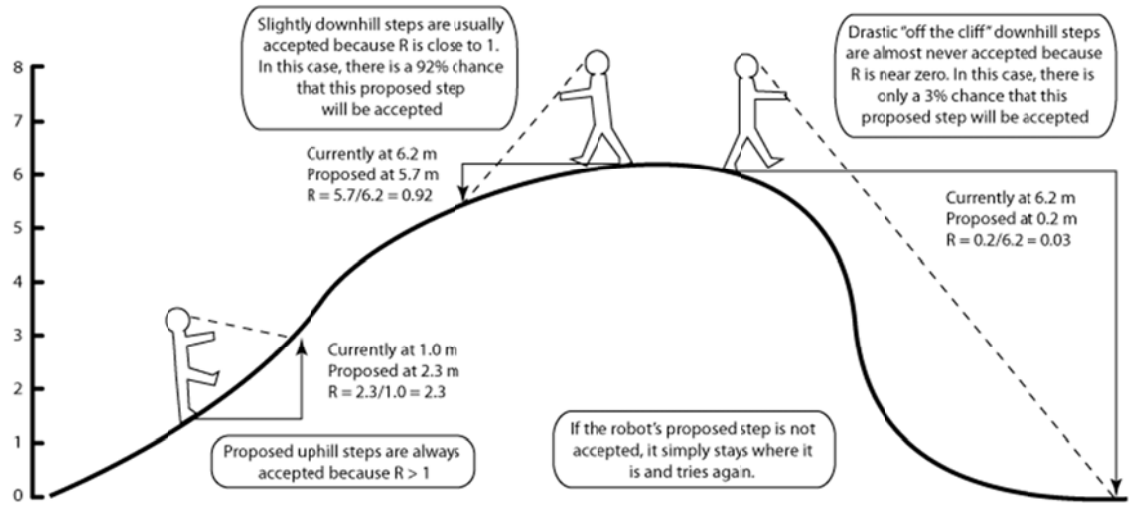


Figure 2.11. The robot metaphor: Principle of the MCMC method. Imagine a robot walking around in a hilly area. From the spot where the robot currently resides a random direction and a random length drawn from a probability distribution is used to propose a step for the robot bringing it to a new position (state). The ratio between the height of the new position and the current position is calculated and termed R. This is equal to the probability for the robot to actually take the proposed step. If R>1 the proposed step is uphill and will always be taken. If R is small the step is major downhill step and hence is very unlikely to be taken. Figure is taken from http://marple.eeb.uconn.edu/mcmcrobot/?page_id=24.

random distance. This new position is only relying on the current position of the robot, which is equal to a Markov chain. The ratio between the height of the new position and the height of the current position is calculated and termed R. This equals the ratio of the likelihood between the present state and a proposed new state. This ratio equals the probability of the robot taking the step. If the ratio R is bigger than 1 the robot will always take the step. Slightly downhill steps will almost always be taken due to a high ratio R resulting in a high probability. This process is repeated for thousands or millions of times. By summarizing the individual states, a posterior distribution around the hill top is created.

In some algorithms Metropolis-coupled MCMC (MC^3) is used. The principle is that several MCMC chains a run in parallel. Sampling only takes place from one chain, called the cold chain. The other chains, termed hot chains, have a probability function which makes steep downhill moves much more likely than the cold chain. The hot chains are hence more likely to "abandon" one hill for a neighboring hill, thereby exploring a much larger area. In each cycle the cold chain can be swapped with one of the hot chains. Whether the swap occurs is based on a probability function. This swapping will in some cases make the cold chain jump valleys and thereby ensure better mixing of the sampled trees by sampling from multiple hills, which equals local optimums.

Another approach for avoiding sampling towards a not global optimum is to start multiple runs with different starting points. If the posterior probability distributions of all runs converge to the same maxima the robustness of the calculated probabilities is strengthened. Probabilistic methods are for many now the golden standard for a wide variety of phylogenetic analysis.

2.6.5 Phylogenetic tree reconciliation

Processes known to contribute to evolution of gene familes is speciation and Incomplete Lineage Sorting (ILS), gene Duplication followed by Loss (DL), and Transfer between coexisting species (T) (Szöllősi et al., 2014). Each of these processes can lead to discord between the gene tree and the corresponding species tree.

Reconciliation is the event of mapping a gene tree to a species tree explaining the evolutionary development of the gene family by the different gene evolution mechanisms. Many algorithms for mapping the gene trees to the species tree exist (Doyon et al., 2011), however they mainly use parsimonious methods for inferring the correct mapping. This involves the usage of event costs for Duplication, Loss and potential Transfer events.

Assigning costs for these events is however not trivial, since no correct costs are known. Small changes in these costs can significantly change the resulting reconciliation.

Probabilistic approaches for the reconciliation of gene trees also exists (Sjöstrand et al., 2012) but until recently these did only consider duplication and loss events. Development of probabilistic methods for reconciliation involving Transfer events is however ongoing and with the DLTRS model by Sjöstrand *et al.* (Sjöstrand et al., 2014) species tree-aware gene trees is constructed in a probabilistic framework. However the current state of the algorithm does not sample the reconciled tree so the inference of specific evolutionary events is not outputted from the algorithm.

2.6.6 Species tree-aware gene trees

Gene trees purely constructed based on molecular data often contains discordance to the related species tree. The discord is due the effect of evolutionary mechanisms such as gene duplication, gene loss, and horizontal gene transfer (HGT), ILS or a combination of these. When lineages diverge each lineage might fix different allelic versions. The pattern of fixing alleles might not be in concordance with the species tree and can result in discord, which is known as ILS. Hence ILS is usually most prevalent when branch lengths are short, i.e. when dealing with closely related species. It is important that algorithms can distinguish between ILS and gene duplication and loss.

There exist numerous algorithms, which take the species tree into account in the gene tree construction. These either uses the species tree information to correct already constructed gene trees, or use the species tree in the construction of the gene trees. These approaches significantly reduce discords between gene trees and species trees by resolving ILS events and at the same time minimizing the parsimonious reconciliation costs. Among the algorithms improving constructed gene trees by using a species tree some of the best are Treefix-DTL (Bansal et al., 2015), exODT (Szöllősi et al., 2013) as well as the probabilistic approach also described above, PRiME-DLTRS (Sjöstrand et al., 2014).

2.7 References

Akihisa, T., Tokuda, H., Ukiya, M., Kiyota, A., Yasukawa, K., et al., **2005**. Anti-tumor-initiating effects of monascin, an azaphilonoid pigment from the extract of *Monascus pilosus* fermented rice (red-mold rice). *Chemistry and Biodiversity*, **2**: 1305–1309. doi:10.1002/cbdv.200590101

Bansal, M.S., Wu, Y.-C., Alm, E.J. & Kellis, M., **2015**. Improved gene tree error correction in the presence of horizontal gene transfer. *Bioinformatics*, **31**, 1211–1218. doi:10.1093/bioinformatics/btu806

Bergmann, S., Schümann, J., Scherlach, K., Lange, C., Brakhage, A.A., et al., **2007**. Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nature Chemical Biology*, **3**: 213–217. doi:10.1038/nchembio869

Blanc, P.J., Loret, M.O., Santerre, a L., Pareilleux, a, Prome, D., et al., **1994**. Pigments of Monascus. *Journal of Food Science*, **59**: 862–865. doi:10.1111/j.1365-2621.1994.tb08145.x

Brakhage, A.A., **1998**. Molecular regulation of beta-lactam biosynthesis in filamentous fungi. *Microbiology and Molecular Biology Reviews*, **62**: 547–585.

Brakhage, A.A., **2013**. Regulation of fungal secondary metabolism. *Nature Reviews. Microbiology*, **11**: 21–32. doi:10.1038/nrmicro2916

Chen, F.C., Manchard, P.S. & Whalley, W.B., **1969**. The structure of monascin. *Journal of the Chemical Society D: Chemical Communications*, **1969**: 130. doi:10.1039/c29690000130

Chen, W., Xie, T., Shao, Y. & Chen, F., **2012**. Genomic characteristics comparisons of 12 food-related filamentous fungi in tRNA gene set, codon usage and amino acid composition. *Gene*, **497**: 116–124. doi:10.1016/j.gene.2012.01.016

Chiang, Y.M., Szewczyk, E., Davidson, A.D., Keller, N., Oakley, B.R., et al., **2009**. A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *Journal of the American Chemical Society*, **131**: 2965–2970. doi:10.1021/ja8088185

Choe, D., Lee, J., Woo, S. & Shin, C.S., **2012**. Evaluation of the amine derivatives of *Monascus* pigment with anti-obesity activities. *Food Chemistry*, **134**: 315–323. doi:10.1016/j.foodchem.2012.02.149

Cox, R.J., **2007**. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Organic and Biomolecular Chemistry*, **5**: 2010–2026. doi:10.1039/b704420h

Cox, R.J. & Simpson, T.J., **2009**. Fungal type I polyketide synthases. *Methods in Enzymology*, **459**: 49-78. doi:10.1016/S0076-6879(09)04603-5

Crawford, J.M., Dancy, B.C.R., Hill, E. a, Udwary, D.W. & Townsend, C. a, **2006**. Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase. *Proceedings of the National Academy of Sciences of the United States of America*, **103**: 16728–16733. doi:10.1073/pnas.0604112103

Crawford, J.M. & Townsend, C.A., **2010**. New insights into the formation of fungal aromatic polyketides. *Nature Reviews. Microbiology*, **8**: 879–889. doi:10.1038/nrmicro2465

Domínguez-Espinosa, R.M. & Webb, C., **2003**. Submerged fermentation in wheat substrates for production of *Monascus* pigments. *World Journal of Microbiology and Biotechnology*, **19**: 329–336. doi:10.1023/A:1023609427750

Doyon, J.P., Ranwez, V., Daubin, V. & Berry, V., **2011**. Models, algorithms and programs for phylogeny reconciliation. *Briefings in Bioinformatics*, **12**: 392–400. doi:10.1093/bib/bbr045

Du, L. & Lou, L., **2010**. PKS and NRPS release mechanisms. *Natural Product Reports*, **27**: 255–278. doi:10.1039/b912037h

Dufossé, L., Galaup, P., Yaron, A., Arad, S.M., Blanc, P., et al., **2005**. Microorganisms and microalgae as sources of pigments for food use: A scientific oddity or an industrial reality? *Trends in Food Science and Technology*, **16**: 389–406. doi:10.1016/j.tifs.2005.02.006

Felsenstein, J., **2004**. *Inferring Phylogenies*, Sunderland, Massachusetts: Sinauer Associates, Inc.

Feng, Y., Shao, Y. & Chen, F., **2012**. *Monascus* pigments. *Applied Microbiology and Biotechnology*, **96**: 1421–1440.

Fielding, B.C., Haws, E.J., Holker, J.S.E., Powell, A.D.G., Robertson, A., et al., **1960**. Monascorubrin. *Tetrahedron Letters*, **1**: 24–27. doi:10.1016/S0040-4039(01)82691-5

Fielding, B.C., Holker, J.S.E., Jones, D.F., Powell, A.D.G., Richmond, K.W., et al., **1961**. 898. The chemistry of fungi. Part XXXIX. The structure of monascin. *Journal of the Chemical Society (Resumed)*, **39**: 4579-4589. doi:10.1039/jr9610004579

Finkelstein, E., Amichai, B. & Grunwald, M.H., **1996**. Griseofulvin and its uses. *International Journal of Antimicrobial Agents*, **6**: 189–194. doi:10.1016/0924-8579(95)00037-2

Fleming, A., **1929**. On the antibacterial action of cultures of a Penicillium, with special reference to their use in the isolation of *B. influenzae*. *British Journal of Experimental Pathology*, **10**: 226–236.

Gao, J.-M., Yang, S.-X. & Qin, J.-C., **2013a**. Azaphilones: chemistry and biology. *Chemical Reviews*, **113**: 4755–4811. doi:10.1021/cr300402y

Gunasekaran, S. & Poorniammal, R., **2008**. Optimization of fermentation conditions for red pigment production from *Penicillium sp.* under submerged cultivation. *African Journal of Biotechnology*, **7**: 1894–1898.

Hajjaj, H., François, J.-M., Goma, G. & Blanc, P.J., **2012**. Effect of amino acids on red pigments and citrinin production in *Monascus ruber*. *Journal of Food Science*, **77**: M156–M159. doi:10.1111/j.1750-3841.2011.02579.x

Hajjaj, H., Klaebe, A., Loret, M.O., Tzedakis, T., Goma, G., et al., **1997**. Production and identification of N-glucosylrubropunctamine and N- glucosylmonascorubramine from *Monascus ruber* and occurrence of electron donor-acceptor complexes in these red pigments. *Applied and Environmental Microbiology*, **63**: 2671–2678.

Hashimoto, M., Nonaka, T. & Fujii, I., **2014**. Fungal type III polyketide synthases. *Natural Product Reports*, **31**: 1306–1317. doi:10.1039/c4np00096j

Hawksworth, D.L. & Rossman, a Y., **1997**. Where are all the undescribed fungi? *Phytopathology*, **87**: 888–891. doi:10.1094/PHYTO.1997.87.9.888

Haws, E.J. & Holker, J.S.E., **1961**. 742. The chemistry of fungi. Part XXXVIII. Further evidence for the structure of rubropunctatin. *Journal of the Chemical Society (Resumed)*, **IV**: 3820-3822. doi:10.1039/jr9610003820

Haws, E.J., Holker, J.S.E., Kelly, A., Powell, A.D.G. & Robertson, A., **1959**. 722. The chemistry of fungi. Part XXXVII. The structure of rubropunctatin. *Journal of the Chemical Society (Resumed)*, 3598-3610. doi:10.1039/jr9590003598

Hertweck, C., **2009**. The biosynthetic logic of polyketide diversity. *Angewandte Chemie (International ed. in English)*, **48**: 4688–4716. doi:10.1002/anie.200806121

Hibbett, D.S., Binder, M., Bischoff, J.F., Blackwell, M., Cannon, P.F., et al., **2007**. A higher-level phylogenetic classification of the Fungi. *Mycological Research*, **111**: 509–547. doi:10.1016/j.mycres.2007.03.004

Hoffmeister, D. & Keller, N.P., **2007**. Natural products of filamentous fungi: enzymes, genes, and their regulation. *Natural Product Reports*, **24**: 393–416. doi:10.1039/b603084j

Hsieh, Y.-H.P., Pao, S. & Li, J., **2008**. Traditional chinese fermented foods. In E. R. Farnworth, ed. *Handbook of Fermented Functional Foods*. Boca Raton: CRC Press, pp. 433–474.

Huang, H.T., **2000**. Science and Civilisation in China: Volume 6, Biology and Biological Technology, Part 5, Fermentations and Food Science. Joseph Needham series. In *Science and Civilisation in China*. Cambridge, England: Cambridge University Press, p. 193.

Inouye, Y., Nakanishi, K., Nishikawa, H., Ohashi, M., Terahara, A., et al., **1962**. Structure of monascoflavin. *Tetrahedron*, **18**: 1195–1203. doi:10.1016/S0040-4020(01)99287-7

Jongrungruangchok, S., Kittakoop, P., Yongsmith, B., Bavovada, R., Tanasupawat, S., et al., **2004**. Azaphilone pigments from a yellow mutant of the fungus *Monascus kaoliang*. *Phytochemistry*, **65**: 2569–2575. doi:10.1016/j.phytochem.2004.08.032

Jung, H., Kim, C., Kim, K. & Shin, C.S., **2003**. Color characteristics of monascus pigments derived by fermentation with various amino acids. *Journal of agricultural and food chemistry*, **51**: 1302–6. doi:10.1021/jf0209387

Jung, H., Kim, C. & Shin, C.S., **2005**. Enhanced photostability of monascus pigments derived with various amino acids via fermentation. *Journal of agricultural and food chemistry*, **53**: 7108–14. doi:10.1021/jf0510283

Jůzlová, P., Martínková, L. & Křen, V., **1996**. Secondary metabolites of the fungus *Monascus*: A review. *Journal of Industrial Microbiology*, **16**: 163–170. doi:10.1007/BF01569999

Keller, N.P. & Hohn, T.M., **1997**. Metabolic pathway gene clusters in filamentous fungi. *Fungal Genetics and Biology*, **21**: 17–29. doi:10.1006/fgbi.1997.0970

Keller, N.P., Turner, G. & Bennett, J.W., **2005**. Fungal secondary metabolism - from biochemistry to genomics. *Nature Reviews. Microbiology*, **3**: 937–947. doi:10.1038/nrmicro1286

Kliebenstein, D.J. & Osbourn, A., **2012**. Making new molecules - evolution of pathways for novel metabolites in plants. *Current Opinion in Plant Biology*, **15**: 415–423. doi:10.1016/j.pbi.2012.05.005

Kumasaki, S., Nakanishi, K., Nishikawa, E. & Ohashi, M., **1962**. Structure of monascorubrin. *Tetrahedron*, **18**: 1171–1184. doi:10.1016/S0040-4020(01)99285-3

Lee, B., Park, N., Piao, H.Y. & Chung, W., **2001**. Production of red pigments by *Monascus purpureus* in submerged culture. *Biotechnology and Bioprocess Engineering*, **6**: 341–346. doi: 10.1007/BF02933003

Lin, T.F., Yakushijin, K., Buchi, G.H. & Demain, A. L., **1992**. Formation of water-soluble *Monascus* red pigments by biological and semi-synthetic processes. *Journal of Industrial Microbiology*, **9**: 173–179. doi:10.1007/BF01569621

Manchand, P.S., Whalley, W.B. & Chen, F., **1973**. Isolation and structure of ankaflavin: A new pigment from *Monascus anka*. *Phytochemistry*, **12**: 2531–2532. doi:10.1016/0031-9422(73)80470-4

Mapari, S.A.S., Hansen, M.E., Meyer, A.S. & Thrane, U., **2008**. Computerized screening for novel producers of *Monascus*-like food pigments in *Penicillium* species. *Journal of Agricultural and Food Chemistry*, **56**: 9981–9. doi:10.1021/jf801817q

Mapari, S.A.S., Meyer, A.S., Frisvad, J.C. & Thrane, U., **2009a**. Production of *Monascus*-like azaphilone pigments. Patent No. WO2009/026923 A2

Mapari, S.A.S., Meyer, A.S. & Thrane, U., **2009b**. Photostability of natural orange-red and yellow fungal pigments in liquid food model systems. *Journal of Agricultural and Food Chemistry*, **57**: 6253–61. doi:10.1021/jf900113q

Martínková, L., Jůzlová, P. & Veselý, D., **1995**. Biological activity of polyketide pigments produced by the fungus *Monascus*. *Journal of Applied Bacteriology*, **79**: 609–616. doi:10.1111/j.1365-2672.1995.tb00944.x

Matsuzaki, K., Ikeda, H., Masuma, R., Tanaka, H. & Omura, S., **1995**. Isochromophilones I and II, novel inhibitors against gp120-CD4 binding produced by *Penicillium multicolor* FO-2338. I. Screening, taxonomy, fermentation, isolation and biological activity. *The Journal of Antibiotics*, **48**: 703–707.

McDaniel, R., Ebert-khosla, S., Fu, H., Hopwoodt, D.A. & Khosla, C., **1994**. Engineered biosynthesis of novel polyketides: Influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. *Proceedings of the National Academy of Sciences of the United States of America*, **91**: 11542–11546.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. & Teller, E., **1953**. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, **21**: 1087. doi:10.1063/1.1699114

Musso, L., Dallavalle, S., Merlini, L., Bava, A., Nasini, G., et al., **2010**. Natural and semisynthetic azaphilones as a new scaffold for Hsp90 inhibitors. *Bioorganic and Medicinal Chemistry*, **18**: 6031–6043. doi:10.1016/j.bmc.2010.06.068

Nam, K., Choe, D. & Shin, C.S., **2014**. Antiobesity effect of a jelly food containing the L-tryptophan derivative of *Monascus* pigment in mice. *Journal of Functional Foods*, **9**: 306–314. doi:10.1016/j.jff.2014.05.001

Natsume, M., Takahashi, Y. & Marumo, S., **1985**. (-)-Mitorubrinic acid, a morphogenic substance inducing chlamydospore-like cells, and its related new metabolite, (+)-mitorubrinic acid-B, isolated from *Penicillium funiculosum*. *Agricultural and Biological Chemistry*, **49**: 2517–2519.

Newman, D.J. & Cragg, G.M., **2012**. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *Journal of Natural Products*, **75**: 311–335. doi:10.1021/np200906s

Nielsen, M.T., Klejnstrup, M.L., Rohlfs, M., Anyaogu, D.C., Nielsen, J.B., et al., **2013**. *Aspergillus nidulans* synthesize insect juvenile hormones upon expression of a heterologous regulatory protein and in response to grazing by *Drosophila melanogaster* larvae. *PLoS ONE*, **8**: e73369. doi:10.1371/journal.pone.0073369

Ogihara, J., Kato, J., Oishi, K. & Fujimoto, Y., **2000a**. Biosynthesis of PP-V, a monascorubramine homologue, by *Penicillium* sp. AZ. *Journal of Bioscience and Bioengineering*, **90**: 678–680.

Ogihara, J., Kato, J., Oishi, K. & Fujimoto, Y., **2001**. PP-R, 7-(2-hydroxyethyl)-monascorubramine, a red pigment produced in the mycelia of *Penicillium* sp. AZ. *Journal of Bioscience and Bioengineering*, **91**: 44–47.

Ogihara, J., Kato, J., Oishi, K., Fujimoto, Y. & Eguchi, T., **2000b**. Production and structural analysis of PP-V, a homologue of monascorubramine, produced by a new isolate of *Penicillium* sp. *Journal of Bioscience and Bioengineering*, **90**: 549–554.

Ogihara, J. & Oishi, K., **2002**. Effect of ammonium nitrate on the production of PP-V and monascorubrin homologues by *Penicillium* sp AZ. *Journal of Bioscience and Bioengineering*, **93**: 54–59.

Osmanova, N., Schultze, W. & Ayoub, N., **2010**. Azaphilones: a class of fungal metabolites with diverse biological activities. *Phytochemistry Reviews*, **9**: 315–342. doi:10.1007/s11101-010-9171-3

Pascual-Ahuir, A., Vanacloig-Pedros, E. & Proft, M., **2014**. Toxicity mechanisms of the food contaminant citrinin: Application of a quantitative yeast model. *Nutrients*, **6**: 2077–2087. doi:10.3390/nu6052077

Salomon, H. & Karrer, P., **1932**. Pflanzenfarbstoffe XXXVIII. Ein Farbstoff aus "rotem" Reis, Monascin. *Helvetica Chimica Acta*, **15**: 18–22. doi:10.1002/hlca.19320150104

Samson, R.A., Yilmaz, N., Houbraken, J., Spierenburg, H., Seifert, K.A., et al., **2011**. Phylogeny and nomenclature of the genus *Talaromyces* and taxa accommodated in *Penicillium* subgenus *Biverticillium*. *Studies in Mycology*, **70**: 159–183. doi:10.3114/sim.2011.70.04

Sato, K., Goda, Y., Sakamoto, S.S., Shibata, H., Maitani, T., et al., **1997**. Identification of major pigments containing D-amino acid units in commercial *Monascus* pigments. *Chemical and Pharmaceutical Bulletin*, **45**: 227-229.

Schoch, C.L., Sung, G.-H., López-Giráldez, F., Townsend, J.P., Miadlikowska, J., et al., **2009**. The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. *Systematic Biology*, **58**: 224–239. doi:10.1093/sysbio/syp020

Shao, Y., Lei, M., Mao, Z., Zhou, Y. & Chen, F., **2014**. Insights into *Monascus* biology at the genetic level. *Applied Microbiology and Biotechnology*, **98**: 3911–3922. doi:10.1007/s00253-014-5608-8

Shen, B., **2003**. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology*, **7**: 285–295. doi:10.1016/S1367-5931(03)00020-6

Shurtleff, W. & Aoyagi, A., **2012**. *History of Koji - Grains and/or Soybeans Enrobed with a Mold Culture (300 BCE to 2012)*, Soyinfo Center, Lafayette, CA, USA.

Silveira, S.T., Daroit, D.J. & Brandelli, A., **2008**. Pigment production by *Monascus purpureus* in grape waste using factorial design. *LWT - Food Science and Technology*, **41**: 170–174. doi:10.1016/j.lwt.2007.01.013

Simpson, T.J., **1995**. Polyketide biosynthesis. *Chem. Ind*, 407–411.

Sjöstrand, J., Sennblad, B., Arvestad, L. & Lagergren, J., **2012**. DLRS: Gene tree evolution in light of a species tree. *Bioinformatics*, **28**: 2994–2995. doi:10.1093/bioinformatics/bts548

Sjöstrand, J., Tofigh, A., Daubin, V., Arvestad, L., Sennblad, B., et al., **2014**. A Bayesian method for analyzing lateral gene transfer. *Systematic Biology*, **63**: 409–420. doi:10.1093/sysbio/syu007

Srianta, I., Ristiarini, S., Nugerahani, I., Sen, S.K., Zhang, B.B., et al., **2014**. Recent research and development of *Monascus* fermentation products. *International Food Research Journal*, **21**: 1–12.

Stadler, M., Anke, H., Dekermendjian, K., Reiss, R., Sterner, O., et al., **1995**. Novel bioactive azaphilones from fruit bodies and mycelial cultures of the ascomycete *Bulgaria inquinans* (Fr.). *Natural Product Letters*, **7**: 7–14.

Szöllősi, G.J., Tannier, E., Daubin, V. & Boussau, B., **2015**. The inference of gene trees with species trees. *Systematic Biology*, **64**: e42-e62. doi:10.1093/sysbio/syu048

Szöllősi, G.J., Tannier, E., Lartillot, N. & Daubin, V., **2013**. Lateral gene transfer from the dead. *Systematic Biology*, **62**: 386–397. doi:10.1093/sysbio/syt003

Tobert, J.A., **2003**. Lovastatin and beyond: the history of the HMG-CoA reductase inhibitors. *Nature Reviews. Drug Discovery*, **2**: 517–526. doi:10.1038/nrd1112

Went, C., **1895**. *Monascus purpureus*, Le Champignon de l'ang-quac. *Annales des Sciences Naturelles*, **Ser. 8**: 1–18.

Whalley, W.B., **1963**. The sclerotiorin group of fungal metabolites: their structure and biosynthesis. *Pure and Applied Chemistry*, **7**: 565–587. doi:10.1351/pac196307040565

Wisecaver, J.H. & Rokas, A., **2015**. Fungal metabolic gene clusters: caravans traveling across genomes and environments. *Frontiers in Microbiology*, **6**: 1–11. doi:10.3389/fmicb.2015.00161

Yang, Y., Liu, B., Du, X., Li, P., Liang, B., et al., **2015**. Complete genome sequence and transcriptomics analyses reveal pigment biosynthesis and regulatory mechanisms in an industrial strain, *Monascus purpureus* YY-1. *Scientific Reports*, **5**: 8331. doi:10.1038/srep08331

Yazar, S. & Omurtag, G.Z., **2008**. Fumonisins, trichothecenes and zearalenone in cereals. *International Journal of Molecular Sciences*, **9**: 2062–2090. doi:10.3390/ijms9112062

Zabala, A.O., Xu, W., Chooi, Y.-H. & Tang, Y., **2012**. Characterization of a silent azaphilone gene cluster from *Aspergillus niger* ATCC 1015 reveals a hydroxylation-mediated pyran-ring formation. *Chemistry and biology*, **19**: 1049–1059. doi:10.1016/j.chembiol.2012.07.004

Zheng, Y., Xin, Y., Shi, X. & Guo, Y., **2010**. Cytotoxicity of *Monascus* pigments and their derivatives to human cancer cells. *Journal of Agricultural and Food Chemistry*, **58**: 9523–9528. doi:10.1021/jf102128t

Chapter 3

Talaromyces atroroseus, a new species efficiently producing industrially relevant red pigments

This chapter presents the paper "*Talaromyces atroroseus*, a new species efficiently producing industrially relevant red pigments" published in PLOS ONE | December 2013 | Volume 8 | Issue 12. My primary contribution to the work presented is the extrolite profiling of *Talaromyces atroroseus* strains, especially adding LC/MS support for the identification of secondary metabolites.

For the readability of the entire PhD thesis the article has been adapted to the thesis format. Text, figures and tables presented are identical to the content of the published paper, which can be found in Appendix A together with the supporting information for the paper.

Talaromyces atroroseus, a new species efficiently producing industrially relevant red pigments

Jens C. Frisvad[1*], Neriman Yilmaz[2, 3], Ulf Thrane[1], Kasper Bøwig Rasmussen[1], Jos Houbraken[2], Robert A. Samson[2]

[1]Department of Systems Biology, Technical University of Denmark, Søltofts Plads B221, DK-2800 Kgs. Lyngby, Denmark
[2]CBS–KNAW Fungal Biodiversity Centre, Uppsalalaan 8, NL-3584 CT Utrecht, the Netherlands
[3]Microbiology, Department of Biology, Utrecht University, Padualaan 8, NL-3584 CH Utrecht, the Netherlands.

Keywords: red pigment, *Penicillium*, azaphilones, mitorubrins, *Monascus* pigments

PLoS ONE 8(12): e84102. doi:10.1371/journal.pone.0084102

Editor: Scott E. Baker, Pacific Northwest National Laboratory, United States of America

Received July 11, 2013; **Accepted** November 8, 2013; **Published** December 19, 2013

Funding: Part of this work was supported by the Danish Research Agency for Technology and Production Grant 09-064967 and an equipment grant from Agilent Technologies. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: A commercial funder (Agilent Technologies) has provided some instruments via an Agilent Thought Leader Award to Jens C. Frisvad. This commercial funder (along with any other relevant declarations relating to employment, consultancy, patents, products in development or marketed products etc.) does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

[*]E-mail: jcf@bio.dtu.dk

3.1 Abstract

Some species of *Talaromyces* secrete large amounts of red pigments. Literature has linked this character to species such as *Talaromyces purpurogenus*, *T. albobiverticillius*, *T. marneffei* and *T. minioluteus* often under earlier *Penicillium* names. Isolates identified as *T. purpurogenus* have been reported to be interesting industrially and they can produce extracellular enzymes and red pigments, but they can also produce mycotoxins such as rubratoxin A and B and luteoskyrin. Production of mycotoxins limits the use of isolates of a particular species in biotechnology. *Talaromyces atroroseus* sp. nov., described in this study, produces the azaphilone biosynthetic families mitorubrins and *Monascus* pigments without any production of mycotoxins. Within the red pigment producing clade, *T. atroroseus* resolved in a distinct clade separate from all the other species in multigene phylogenies (ITS, β-tubulin and RPB1), which confirm its unique nature. *Talaromyces atroroseus* resembles *T. purpurogenus* and *T. albobiverticillius* in producing red diffusible pigments, but differs from the latter two species by the production of glauconic acid, purpuride and ZG–1494α and by the dull to dark green, thick walled ellipsoidal conidia produced. The type strain of *Talaromyces atroroseus* is CBS 133442

3.2 Introduction

Monascus species are known to produce six major azaphilone pigments being the yellow monascin and ankaflavin; the orange monascorubrin and rubropunctatin and the red monascorubramine and rubropunctamine, in addition to more than 20 related pigments (Feng et al., 2012; Patakova, 2013). Another azaphilone series of yellow pigments is even more widespread in *Talaromyces*, i.e. the mitorubrins (Samson et al., 1989; Van Reenen-Hoekstra et al., 1990; Frisvad et al., 1990). The red pigment producer *Monascus purpureus* has been used primarily in Southern China, Japan and Southeast Asia for making red rice wine, red soybean cheese and Anka (red rice) (Lin et al., 2008). A problem is that some samples of *Monascus*–fermented rice have been found to contain the mycotoxin citrinin (Liu et al., 2005), but also that *Monascus* isolates also often produce mevinolin, a drug that is also unwanted in foods (Patakova, 2013). The production of such mycotoxins and drugs limits the use of *Monascus* for industrial purposes, but since citrinin has not been found in any *Talaromyces* species, the latter may be a good alternative for red pigment production. Studies have shown that polyketide azaphilone *Monascus* red pigments and/or their amino acid derivatives are naturally produced

by *Talaromyces aculeatus*, *T. pinophilus*, *T. purpurogenus* and *T. funiculosus* (Mapari et al., 2008; Mapari et al., 2009). *Talaromyces amestolkiae*, *T. ruber* and *T. stollii* also produce azaphilone polyketides, as recently described by Yilmaz *et al.* (Yilmaz et al., 2012), but in those three species the pigment are not diffusing into the growth medium. *Talaromyces amestolkiae* and *T. stollii* were isolated from immuno-compromised patients and are potential human pathogens, while *T. purpurogenus* produces mycotoxins such as rubratoxins A and B, rugulovasins, and luteoskyrin (Yilmaz et al., 2012). These factors limit the use of these species for biotechnological production of azaphilone pigments. In the current study we describe a new *Talaromyces* species, *T. atroroseus*, which secretes large amounts of *Monascus* red pigments, without the production of any known mycotoxins.

3.3 Materials and Methods

3.3.1 Strains

Cultures were obtained from the CBS-KNAW Fungal Biodiversity Centre culture collection, Utrecht, the Netherlands. Fresh isolates deposited in the working collection of the Department of Applied and Industrial Mycology (DTO) housed at CBS, and strains from the IBT collection at DTU Systems Biology in Kgs. Lyngby, Denmark were also included in this study. Strains are listed in Table 3.1. KAS strain numbers are from the fungal collection of Keith A. Seifert, Ottawa, Canada.

3.3.2 Morphological analysis

Macroscopic characters were studied on agar media Czapek-Dox yeast autolysate agar (CYA), CYA supplemented with 5 % NaCl (CYAS), yeast extract sucrose agar (YES), creatine sucrose agar (CREA), dichloran 18 % glycerol agar (DG18), oatmeal agar (OA) and malt extract agar (Oxoid) (MEA). The isolates were also tested on CYA at 37 °C and on Blakeslee malt extract agar (MEA2). All media were prepared as described by Samson *et al.* (Samson et al., 2010). The strains were inoculated in three points onto media in 90-mm Petri dishes and incubated for 7 d at 25 °C in darkness. After incubation, the colony diameters on the various agar media were measured. Colonies were photographed with a Canon EOS 400D. Species were characterized microscopically by preparing slides from MEA. Lactic acid was used as mounting fluid. Specimens were examined using a Zeiss AxioSkop2 plus microscope.

Table 3.1. Strains used in this study of *Talaromyces atroroseus* and related species.

CBS No.	Other Collection No.	Species	Information and Origin
206.89	IFO 6580, IBT 3960; DTO 41F4	*T. albobiverticillius*	Unknown, Japan
238.95	IBT 11181, CBS 123796	*T. atroroseus*	Red sweet bell pepper, Kgs. Lyngby, Denmark
234.60	DTO 37A4	*T. atroroseus*	Unknown, Germany
257.37	DTO 37A3	*T. atroroseus*	Ex air in nitrite factory, Germany
313.63	DTO 41G2	*T. atroroseus*	*Vitis vinifera* fruit, South Africa
364.48	ATCC 9777, IMI 040037, NRRL 1061, QM 6760, DTO 178A3, IBT 4458, IBT 11180	*T. atroroseus*	Unknown, Darien, Manchuria, China
391.96	DTO 41G8	*T. atroroseus*	Unknown, Tanzania
113139	IBT 3967, NRRL 1147, DTO 177I2	*T. atroroseus*	Unknown, USA
113167	DTO 39I2, DTO 39I3	*T. albobiverticillius*	Unknown, unknown
113168	IBT 31347, DTO 39H9, DTO 177I9	*T. albobiverticillius*	Sputum of patient, male, Copenhagen, Denmark
113153	IBT 3458, NRRL 1136, DTO 37A7	*T. atroroseus*	Ex mixed culture, Arlington Farm, Virginia, USA
124294	IBT 23082	*T. atroroseus*	Tropical rainforest, Peru
133440	BCRC 34774, DTO 166E5, IBT 31667	Type of *T. albobiverticillius*	Decaying leaves of a broad–leaved tree, Taiwan
133441	BCRC 34775, DTO 166E6, IBT 31668	*T. albobiverticillius*	Decaying leaves of a broad–leaved tree, Taiwan
133442	KAS 3778, DTO 178A4, IBT 32470	Type of *T. atroroseus*	House dust, South Africa
133443	IBT 29388, DTO 189D4	*T. atroroseus*	Contamination in petri dish, Lyngby, Denmark
133444	IMI 163167, IBT 23702, DTO 189C2	*T. albobiverticillius*	*Punica granata*, unknown
133447	DTO 81I2	*T. atroroseus*	Swab sample from cheese warehouse, the Netherlands
133448	DTO 157G5	*T. albobiverticillius*	Pomegranate, Turkey
133449	IBT 29464, DTO 189D5	*T. atroroseus*	Mouse dung, Høve Strand, Denmark
133450	FRR 75, IBT 4454, DTO 188I9	*T. atroroseus*	Soil Murrumbidgee irrigation Area, New South Wales, Australia
133452	NRRL 2120, IBT 3547, DTO 193H9	*T. albobiverticillius*	Cotton duck, Panama
113154	R.B., IMI 090178, NRRL 1214, IBT 3645, IBT 4428, CBS 127571	*T. atroroseus*	"Parasite" in *Aspergillus niger* culture, Kansas City, Missouri, USA
	TA85S-28-H2, AZ, IAM 15392, JCM 23216, IBT 32650	*T. atroroseus*	Soil, Thailand
	IBT 20955	*T. atroroseus*	Air root in white mangrove, Can de Aruca, Paria Bay, Venezuela
	IBT 4466	*T. albobiverticillius*	*Punica granata*, imported to Denmark

3.3.3 DNA extraction, PCR amplification and sequencing

Strains were grown for 7 to 14 d on MEA prior to DNA extraction. DNA was extracted using the UltracleanTM Microbial DNA isolation Kit (MoBio, Solana Beach, U.S.A.). The extracted DNA was stored at -20 °C. The ITS regions, regions of the β-tubulin and *RPB1* genes were amplified and sequenced according to methods previously described (Houbraken et al., 2007; Houbraken et al., 2011; Houbraken & Samson, 2011; Samson et al., 2011).

3.3.4 Data analysis

Sequence contigs were assembled using Seqman from DNAStar Inc. Newly generated ITS, β- tubulin and *RPB1* sequences were included in a data set obtained from the Samson *et al.* (Samson et al., 2011) study. Data sets were aligned using Muscle software within MEGA5 (Tamura et al., 2011). Neighbour–joining analysis on the individual data sets was performed in MEGA5 and confidence in nodes determined using bootstrap analysis with 1000 replicates. *Talaromyces galapagensis* (CBS 751.74^{T}) was selected as a suitable out-group in all the phylogenies. The newly generated sequences were deposited in GenBank (accession numbers, see Table 3.1 and Figure 3.1– 3.3).

3.3.5 Extrolites

Cultures grown on CYA and YES for 7 d at 25 °C were used for extrolite extractions. Extracts were analysed by HPLC using alkylphenone retention indices and diode array UV–VIS detection as described by (Frisvad & Thrane, 1987; Nielsen et al., 2011; Houbraken et al., 2012), using three 6 mm agar plugs. Standards of extrolites from the collection at DTU Systems Biology (Denmark) were used to compare the extrolites from the species under study (Nielsen et al., 2011). The extrolite extractions from *T. atroroseus* CBS 133450, CBS 113154 and CBS 123796 were also analysed by ultra high performance liquid chromatography high-resolution mass spectrometry (UHPLC-HRMS). Liquid chromatography was performed on an Agilent 1290 Infinity LC system with a DAD-detector coupled to an Agilent 6550 iFunnel Q-TOF with an electrospray ionization source. The separation was performed on a 2.1 x 250 mm, 2.7 μm Poroshell 120 Phenyl-Hexyl column (Agilent) at 60 oC with a water-acetonitrile gradient (both with 20 mM formic acid) going from 10 % (vol/vol) to 100 % acetonitrile in 15 min followed by 2.5 min with 100 % acetonitrile and then returning to the start conditions for 2.5 min for equilibration before next sample. All time the flow rate was kept at 0.35 mL/min. HRMS was performed in ESI+ and extrolites were identified with targeted search on accurate

mass of $[M+H]^+$ and $[M+Na]^+$ using Agilent MassHunter Qualitative Analysis B.06.00 software and a database of potential extrolites in *T. atroroseus* with support from UV-VIS spectra. The list of compounds searched for including the extrolite standards can be found in appendix Table A.1.

3.3.6 Nomenclature

1. The electronic version of this article in Portable Document Format (PDF) in a work with an ISSN or ISBN will represent a published work according to the International Code of Nomenclature of algae, fungi, and plants, and hence the new names contained in the electronic publication of a PLOS ONE article are effectively published under that Code from the electronic edition alone, so there is no longer any need to provide printed copies. In addition, new names contained in this work have been submitted to MycoBank from where they will be made available to the Global Name Index. The unique MycoBank number can be resolved and the associated information viewed through any standard web browser by appending the MycoBank number contained in this publication to the prefix http://www.mycobank.org/MB. The online version of this work is archived and available from the following digital repositories. PubMed Central, LOCKSS.

2. Repository of ***Talaromyces atroroseus*** Yilmaz, Frisvad, Houbraken & Samson 2013 sp. nov. [urn:lsid:mycobank.org: 804901]

3.4 Results and Discussion

The relationship between the *Talaromyces atroroseus* sp. nov. and its close relatives were studied using multigene phylogenies, based on ITS, *RPB1* and β-tubulin sequences. The aligned datasets were 482, 888 and 374 bp long, respectively. The new species resolved in a clade together with other red pigment producing species such as *T. albobiverticillius*, and *T. minioluteus*. *Talaromyces purpurogenus* resolved in a distantly related clade (Figures 3.1–3.3). Within the red pigment producing clade, *T. atroroseus* resolved in a distinct clade separate from all the other species in all three phylogenies, confirming its unique nature. Historically red pigment production caused a lot of confusion and resulted in numerous misidentifications in literature. This is especially true for *Talaromyces purpurogenus*, *T. ruber*, *Penicillium sanguineum* and *P. crateriforme*. *Penicillium purpurogenum* and *P. rubrum* were described by Stoll (Stoll, 1905). In their

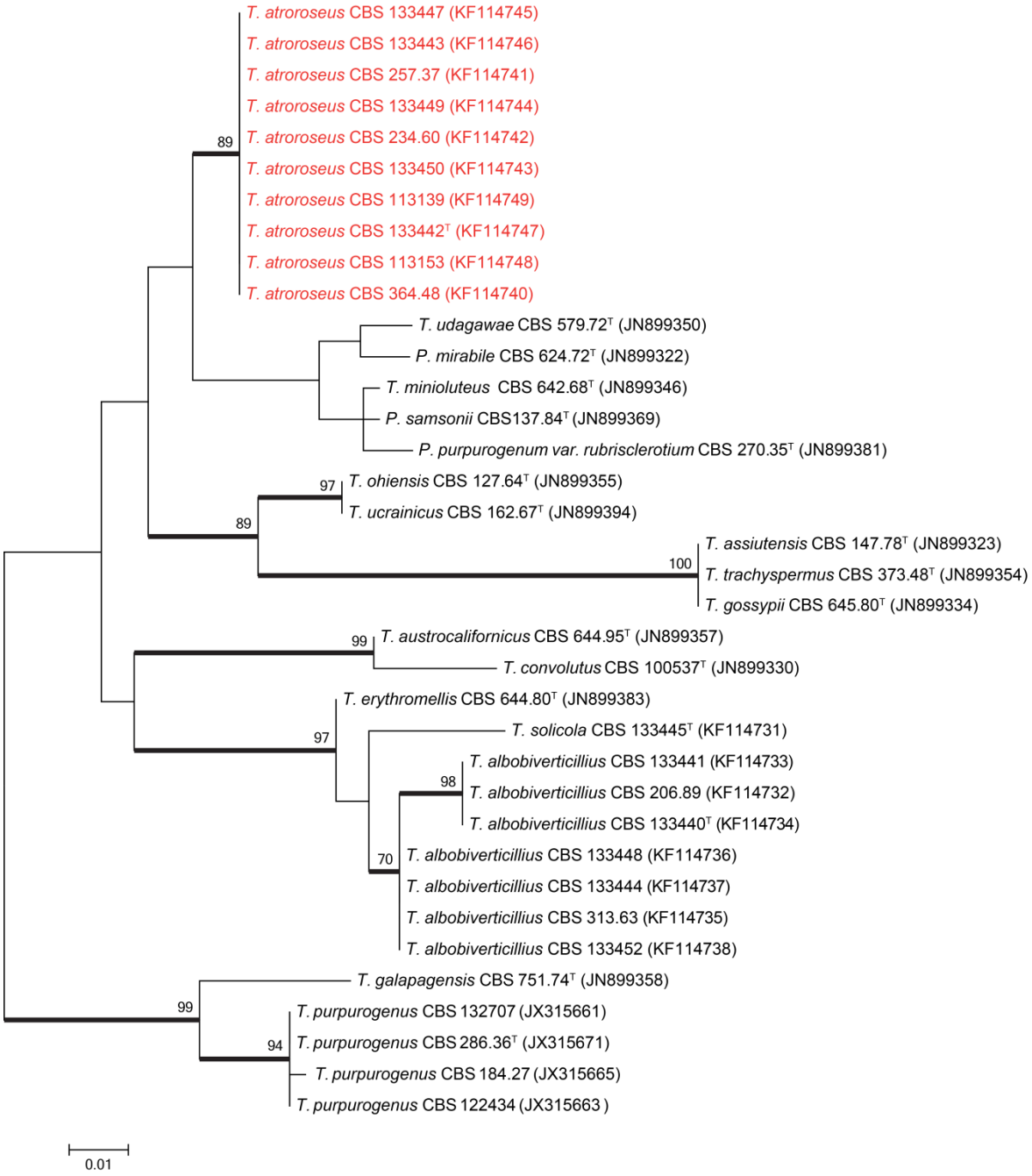


Figure 3.1. Maximum likelihood tree comparing the ITS gene region of *Talaromyces* species closely related to *T. atroroseus.* *Talaromyces galapagensis* and *T. purpurogenus* were used as outgroup. Support in nodes is indicated above thick branches and is represented by bootstrap values higher than 70%. GenBank accession numbers are given between brackets, (T = ex-type). Red coloured names indicate *T. atroroseus* strains.

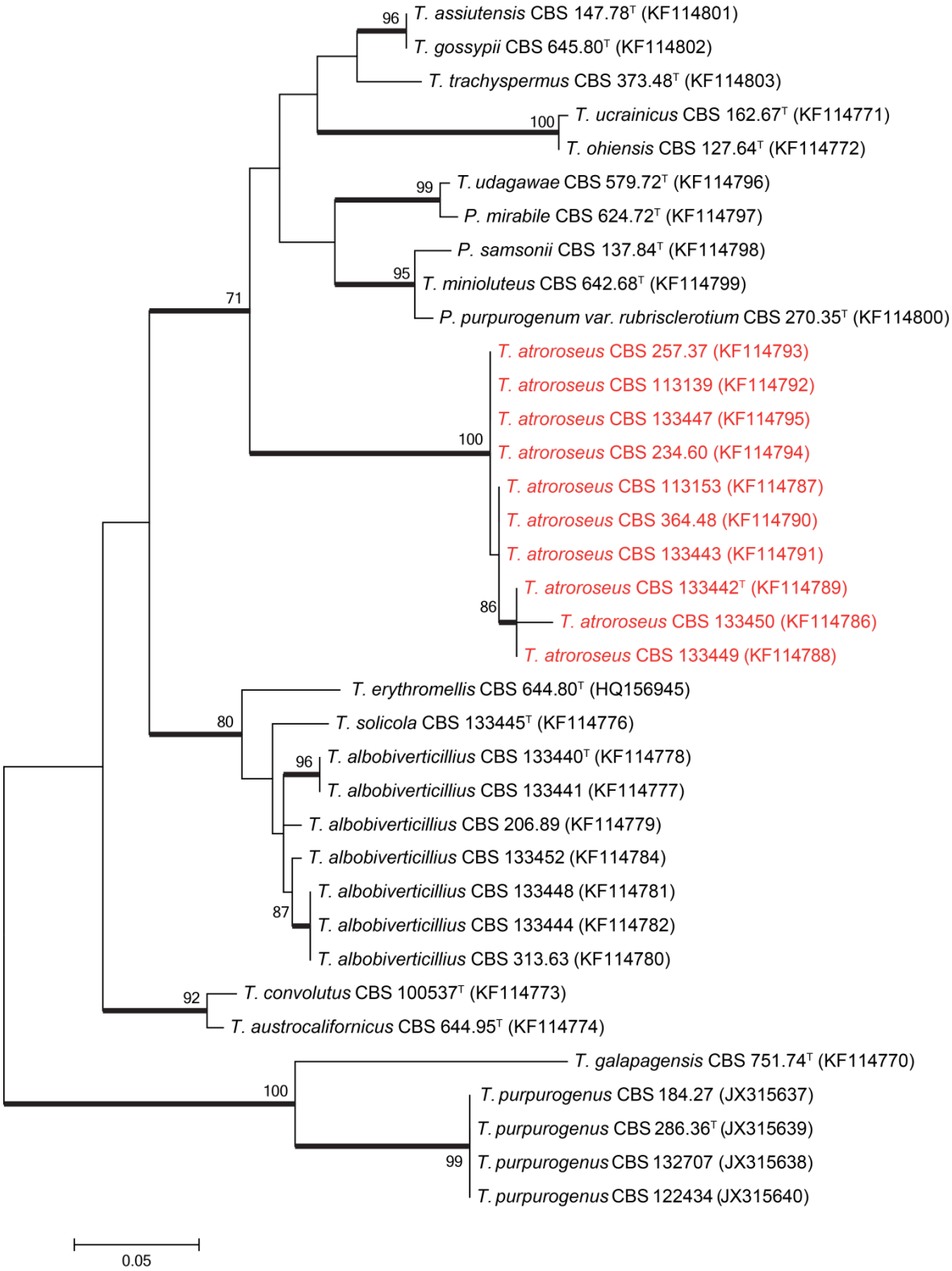


Figure 3.2. Maximum likelihood tree comparing the β–tubulin gene region of *Talaromyces* species closely related to *T. atroroseus.* *Talaromyces galapagensis* and *T. purpurogenus* were used as outgroup. Support in nodes is indicated above thick branches and is represented by bootstrap values higher than 70%. GenBank accession numbers are given between brackets, (T = ex-type). Red coloured names indicate *T. atroroseus* strains.

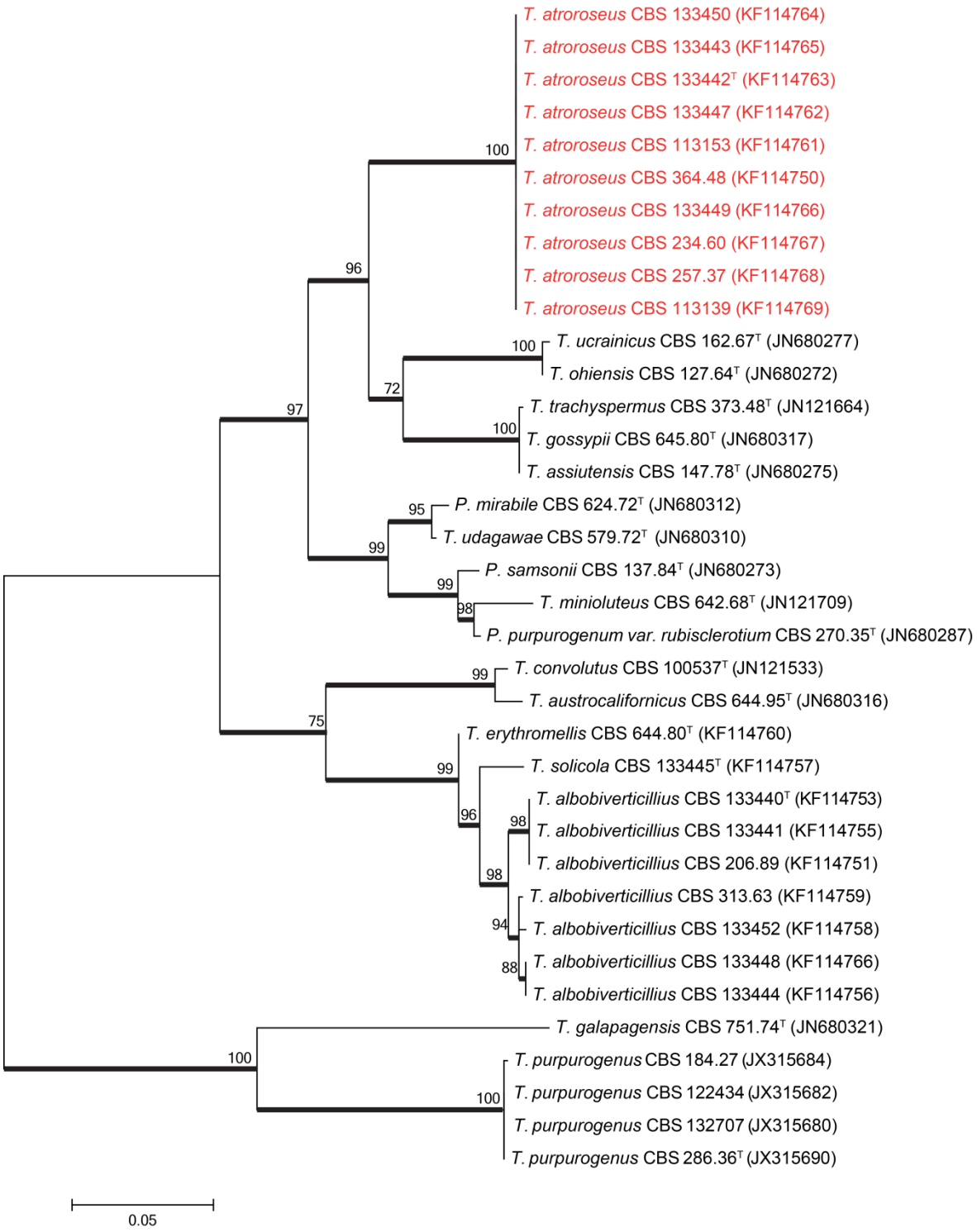


Figure 3.3. Maximum likelihood tree comparing the *RPB1* gene region of *Talaromyces* species closely related to *T. atroroseus*. *Talaromyces galapagensis* and *T. purpurogenus* were used as outgroup. Support in nodes is indicated above thick branches and is represented by bootstrap values higher than 70%. GenBank accession numbers are given between brackets, (^{T} = ex-type). Red coloured names indicate *T. atroroseus* strains.

monograph Raper and Thom (Raper & Thom, 1949) also described *P. purpurogenum* and *P. rubrum*. No type material was available for *P. rubrum* therefore Raper and Thom (Raper & Thom, 1949) used two strains to describe *P. rubrum*, NRRL 1062 (= CBS 370.48) and NRRL 2120 (= CBS 133452). Pitt (Pitt, 1980) synonymized *P. rubrum*, *P. crateriforme* and *P. sanguineum* with *P. purpurogenum*. The issues in the *T. purpurogenus* complex were clarified by Yilmaz *et al.* (Yilmaz et al., 2012) who synonymized *Penicillium crateriforme* and *P. sanguineum* with *T. purpurogenus* and they described *T. ruber* as a distinct species. NRRL 1062 remained as *T. ruber* but NRRL 2120 (= CBS 133452) is a different species than *T. ruber*. Our results showed that NRRL 2120 is *T. albobiverticillius*. Raper and Thom (Raper & Thom, 1949) based the *Penicillium purpurogenum* description on NRRL 1061 (= CBS 364.48). However our results show that NRRL 1061 is a typical *T. atroroseus* strain.

Table 3.2. Reported extrolite production by strains verified as *Talaromyces atroroseus* during this study.

Extrolite	Reported producer	Culture collection numbers	References
Glaucanic acid, Glauconic acid	*Penicillium* "R. B.", *P. purpurogenum*	**R.B.** = IMI 090178 = NRRL 1214 = CBS 113154 = IBT 3645 = IBT 4428	(Yuill, 1934; Takashima et al., 1955; Baldwin et al., 1962; Barton et al., 1965a; Barton et al., 1965b; Barton & Sutherland, 1965; Huff et al., 1968; Huff et al., 1972)
N-glutaryl monascorubramine, N-glutaryl rubropunctamine	*P. purpurogenum*	**IBT 11181** = CBS 238.95 = CBS 123796	(Mapari et al., 2009)
N-glutaryl monascorubramine	*P. purpurogenum*	R.B. = IMI 090178 = NRRL 1214 = CBS 113154 = **IBT 3645** = IBT 4428	(Mapari et al., 2009)
Monascorubramine, PP-R	*P. purpurogenum*	**IBT 11180** = CBS 364.48 = ATCC 9777 = IMI 040037 = NRRL 1061 = QM 6760 = IBT 4458	(Mapari et al., 2009)
PP-V, PP-R, PP-O, PP-Y	*Penicillium sp.*	TA85S-28-H2 = **AZ** = IAM 15392 = JCM 23216 = IBT 32650	(Ogihara et al., 2000; Ogihara et al., 2001; Ogihara & Oishi, 2002; Arai et al., 2012; Arai et al., 2013)
Purpuride	*P. purpurogenum*	**CBS 257.37**	(King et al., 1973)
Purpurogenone, Deoxypurpurogenone	*P. purpurogenum*	**CBS 257.37**	(Roberts & Warren, 1955; King et al., 1970; Roberts & Thompson, 1971a; Roberts & Thompson, 1971b)
ZG-1494α	*P. rubrum*	IBT 11181 = **CBS 238.95** = CBS 123796	(West et al., 1996)

Strain numbers in bold are the strain numbers used in the references

Both *Talaromyces purpurogenus* and *T. atroroseus* are common in soil, indoor environments, and fruits. *Talaromyces atroroseus* resembles *T. purpurogenus* and *T. albobiverticillius* in producing red diffusible pigments, but differs from the latter two species by the production of glauconic acid, purpuride and ZG–1494α (Table 3.2 and Figure 3.4) and by the dull to dark green thick walled ellipsoidal conidia produced. Barton *et al.* (Barton et al., 1965a; Barton et al., 1965b) and Barton and Sutherland (Barton & Sutherland, 1965) reported glauconic acid from *P. purpurogenum* IMI 090178, which in the present study has been re-identified as *T. atroroseus*, while ZG–1494α was reported from *P. rubrum* CBS 238.95 (West et al., 1996), which is also a typical *T. atroroseus*. *Talaromyces atroroseus*, *T. purpurogenus* and *T. albobiverticillius* differ from *T. ruber*, *T. amestolkiae* and *T. stollii* by their production of red diffusible pigment. In Table 3.3 many red pigment producers identified as *Penicillium* species are listed that may either be *T. purpurogenus, T. ruber, T. albobiverticillius* or *T. atroroseus*. The strains listed in Table 3.3 were not available for us, so their exact identity cannot be verified.

Many *Talaromyces* species produce striking diffusing red pigments, especially *T. purpurogenus*, *T. atroroseus, T. albobiverticillius, T. minioluteus,* and *T. marneffei.* These red pigments are typically composed of the azaphilone pigments (Figure 3.5) monascorubrin, rubropunctatin, threonine derivative of rubropunctatin, monascorubramine, PP-R (=7–(2–hydroxyethyl)-monascorubramine), rubropunctamine, N-glutarylrubropunctamine, and PP–V (Ogihara et al., 2000; Ogihara et al., 2001; Mapari et al., 2005; Mapari et al., 2008; Mapari et al., 2009). The same family of azaphilones are also known from red rice, where different species of *Monascus* have grown (Feng et al., 2012; Patakova, 2013). These red pigments are of interest for the industry as they are stable and non–toxic and can be used as food colorants (Mapari et al., 2010). The azaphilone pigments can react with amino acids, hence their name, and give intense dark red colours. In addition some of these species produce yellow azaphilone pigments, such as monascin, ankaflavin, monascusone A and B, xanthomonascin A, and another series of yellow mitrorubrin azaphilones: mitorubrin, mitorubrinol, mitorubrinol acetate, mitorubrinic acid, and many other related compounds (Frisvad et al., 1990). Many of these pigments have been reported from or found in *T. atroroseus* in this study (Table 3.2 and Table 3.4). The potential for pigment production has in this study only been investigated in small scale on solid media; however, *T. atroroseus* also produce pigments in liquid cultures under the right conditions (Mapari et al., 2008; Arai et al., 2012). The potential for up scaling the production of red pigments needs to be investigated thoroughly.

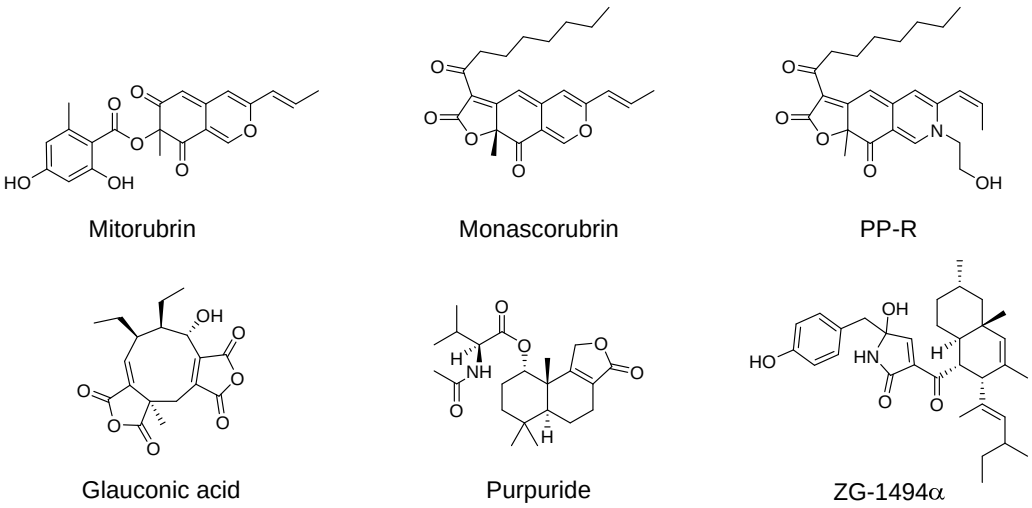


Figure 3.4. Structures of some of the most characteristic compounds produced by *Talaromyces atroroseus*. All six compounds were detected in this study.

Even though sequence variations were observed for *Talaromyces albobiverticillius* strains, morphologically they were similar. Two strains used for the original description of *T. albobiverticillius* were received from Dr. Sung-Yuan Hsieh (Hsieh et al., 2010). These included the type strain CBS 133440^{T} and CBS 133441. These strains were isolated from soil in Taiwan and produce white conidial masses and intense soluble red pigment on various media (Figure 3.6). However, other freshly isolated *T. albobiverticillius* strains produce densely sporulating colonies and do not show any stability for red pigment production. Some of the *Talaromyces albobiverticillius* strains did not produce any soluble pigment such as CBS 133444 and CBS 133448. Strains that did produce red pigments include CBS 113168, and CBS 133452. On MEA only the degraded or mutated strains of *T. albobiverticillius*, such as CBS 133440^{T} and CBS 313.63 produced red pigments. Micromorphologically all *T. albobiverticillius* strains produce long stipes (up to 380 µm) (Figure 3.5). Two strains of *T. albobiverticillius* (CBS 133440T and CBS 133441) have globose to subglobose, smooth conidia; however, the remaining strains produce ellipsoid to fusiform smooth conidia (Figure 3.5). Even though two clades were observed in the phylogenies there are no concordance between observed clades and morphological characters as discussed above. As such, they are considered here as representing one species. Raper and Thom (Raper & Thom, 1949) mentioned a number of colour mutations they observed in strains of *P. citrinum* and *P. chrysogenum*. They stated that colour mutations are encountered as the most common and conspicuous types of mutations,

Table 3.3. Reported extrolite production from strains potentially belonging to *Talaromyces atroroseus*, but not examined during this study.

Extrolite	Reported producer	Strain identifier / Culture collection number	Reference
2,6,7-trihydroxy-3-methyl-naphthalene-1,4-dione	*Penicillium purpurogenum*	JS03-21*	(Wang et al., 2011)
BE-25327	*P. purpurogenum*	F25327 = FERM P-12345	(Kondo et al., 1993)
Dhilirolide A, B, C, D	*P. purpurogenum*	IMI 357108	(de Silva et al., 2011)
Glauconic acid	*P. glaucum*	-*	(Wijkman, 1931)
Gluconic acid	*P. purpurogenum* var. *rubrisclerotium* (= *T. pinophilus*)	No. 2670 = NRRL 1064 = CBS 270.35 = ATCC 4713 = ATCC 52224 = NRRL 1142 = IBT 4302	(Herrick & May, 1928)
(-)-Mitorubrin,	*P. purpurogenum*	JS03-21*	(Wang et al., 2011)
Monascus red pigment	*P.* sp.	HKUCC 8070	(Jiang et al., 2005)
Orsellinic acid	*P. purpurogenum*	JS03-21*	(Wang et al., 2011)
Purpactin A, B, C	*P. purpurogenum*	FO-608 = FERM P-10776	(Tomoda et al., 1991; Nishida et al., 1991)
Purpurester A, B	*P. purpurogenum*	JS03-21*	(Wang et al., 2011)
Purpurquinones A, B, C	*P. purpurogenum*	JS03-21*	(Wang et al., 2011)
Red W59 ($C_{30}H_{34}O_9N_3$)	*P. purpurogenum*	-*	(Pharm Amano, 1974)
Red pigment	*P. purpurogenum*	GH2*	(Mendez-Zavala et al., 2007; Méndez et al., 2011; Espinoza-Hernández et al., 2013)
Red pigment	*P. purpurogenum*	SX01*	(Qiu et al., 2010)
Red pigments	*P. purpurogenum*	DPUA 1275	(Santos-Ebinuma et al., 2013a; Santos-Ebinuma et al., 2013b)
Red pigments	*P. purpurogenum*	-*	(Velmurugan et al., 2010a; Velmurugan et al., 2010b)
Red pigments	*P. sp.*	-*	(Gunasekaran & Poorniammal, 2008)
SL 3238 (C27H41NO7)	*P. purpurogenum*	NRRL 3364	(Bollinger et al., 1972)
TAN-931	*P. purpurogenum*	JS03-21*	(Wang et al., 2011)

Based on the reported morphology and extrolites the strains in the table are by the authors' judgement belonging to *Talaromyces atroroseus* or a closely related species.
* Strain not deposited in any accessible culture collections

Table 3.4. Extrolites of *Talaromyces atroroseus* and *T. albobiverticillius*. Examined by HPLC-DAD and/or UHPLC- HRMS and comparison to standards on the media CYA and YES.

Species	Culture collection number	Extrolites* found
T. atroroseus	CBS 133450[a]	glauconic acid[b], monascorubrin[b], PP-R[b], purpuride[b], purpuroquinone A[b], ZG-1494α[b]
	CBS 113154[a]	glauconic acid, N-glutarylmonascorubramin[b], monascorubrin[b], PP-O[b], PP-R[b], purpuride[b], purpurquinone A[b], ZG-1494α[b]
	CBS 123796[a]	FK17-P2b2[b], glauconic acid, N-glutarylmonascorubramin[b], mitorubrin, mitorubrinol, monascorubrin[b], PP-O[b], PP-R[b], purpuride[b], purpurquinone A[b], purpurogenone, ZG-1494α[b]
	CBS 257.37	monascorubramine, purpuride, several Monascus-red pigments
	CBS 234.60	glauconic acid, monascorubramine, purpuride, ZG-1494α
	CBS 391.96	glauconic acid, monascorubramine, purpuride, ZG-1494α
	CBS 364.48	glauconic acid, monascorubramine, PP-R, purpuride, rubropunctatin, ZG-1494α
	CBS 133447	glauconic acid, purpuride
	CBS 133442	glauconic acid, monascorubramine, purpuride, rubropunctatin
	CBS 113153	glauconic acid, mitorubrin, monascorubramine, monascorubrin, purpuride
	CBS 113139	monascin, monascorubramine
	IBT 3933	glauconic acid, mitorubrin, monascorubramine, a purpactin
	IBT 20955	glauconic acid, monascorubramine, monascorubrin, purpuride, ZG-1494α
	IBT 23082	PP-R (only tested for *Monascus* pigments)
	CBS 133443	glauconic acid, monascorubramine, purpuride
	CBS 133449	glauconic acid, monascorubrin, purpuride
	CBS 313.63	Glauconic acid, mitorubrin, monascorubramine, monascorubrin, purpuride, rubropunctatin
	JCM 23216	glauconic acid, monascorubramine, purpuride
T. albobiverticillius	CBS 113168	mitorubrin, mitorubrinic acid, monascorubramine, PP-R, rubropunctatin, vermicellin
	CBS 313.63	Mitorubrin, monascorubramine, monascorubrin, rubropunctatin
	IBT 4466	mitorubrinic acid, monascorubramine, a purpactin
	CBS 113167	mitorubrin, mitorubrinic acid, monascorubrin, a purpactin
	CBS 133444	mitorubrin, mitorubrinic acid, mitorubrinol
	CBS 133452	mitorubrin, mitorubrinic acid, monascorubramine, rubropunctatin
	CBS 133441	mitorubrin, mitorubrinic acid, monascin, monascorubramine, rubropunctatin, vermicellin

[a] Strains examined by both HPLC-DAD and UHPLC-DAD-HRMS
[b] Extrolites identified by UHPLC-DAD-HRMS
* The extrolites only identified by HPLC-DAD might in some cases not be the actual metabolite but a derivative with the same chromophore and retention on the column

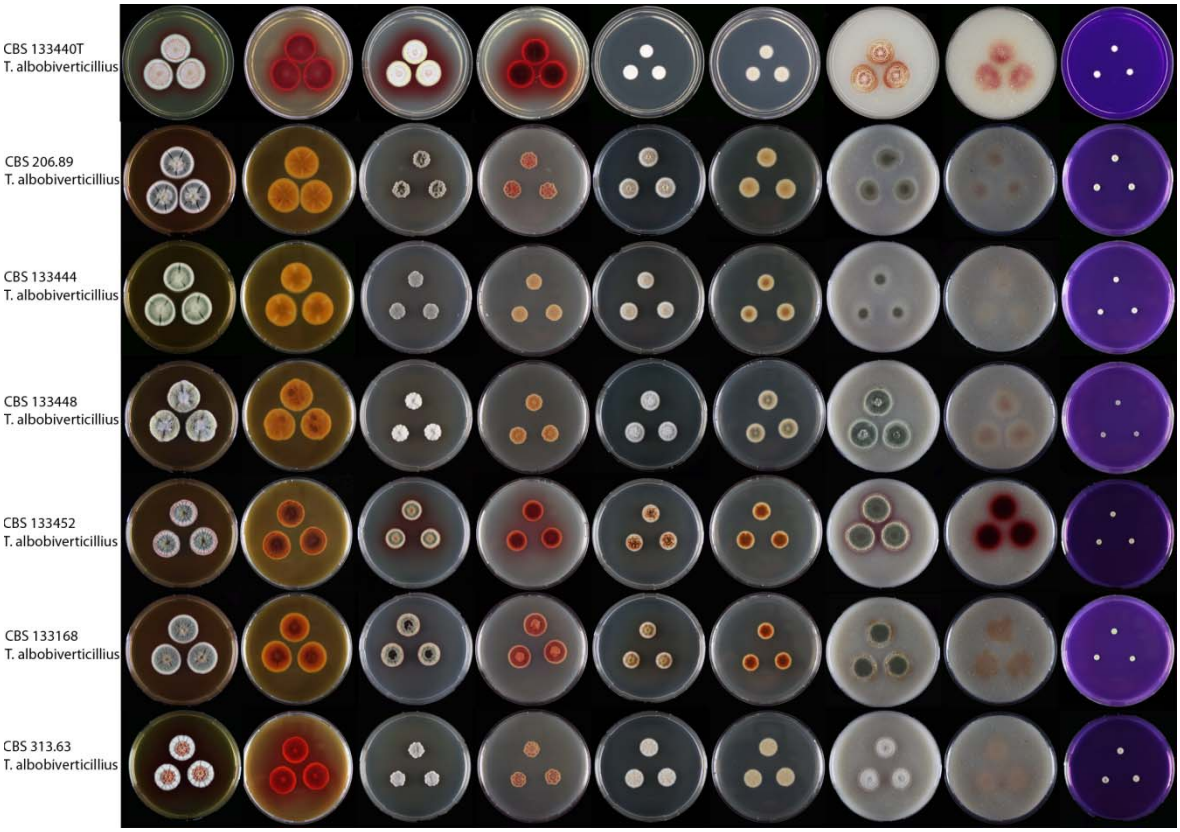


Figure 3.5. Strains of *Talaromyces albobiverticillius* on MEA, CYA, DG18, OA and CREA. Colony obverse and reverse is shown for the first four media.

especially considering mature conidia. Mutations can often be observed when a strain loses its green pigment in its conidia, resulting in a white or tanned colour. Colour mutants are regularly encountered among the strains, which were exposed to artificial stimulations such as ultra-violet, X-ray radiations and neutron bombardment (Raper & Thom, 1949).

Talaromyces atroroseus is considered as the optimal producer of industrially important yellow and red soluble pigments. Another option as a suitable producer of red soluble azaphilone pigments is *T. albobiverticillius*. However *T. albobiverticillius* produces soluble red pigment only in some strains. We speculate that the mitorubrins produced by *Talaromyces atroroseus* are of the (-)-form, as they have been shown to be that for the closely related *Talaromyces purpurogenus* (at that time identified as *Penicillium rubrum*) (Buechi et al., 1965; Marsini et al., 2006). However, Natsume *et al.* (Natsume et al., 1985) and Suzuki *et al.* (Suzuki et al., 1999) found both (+) and (-)-forms in the genus *Talaromyces*, while mitorubrins in *Hypoxylon* and other related genera are of the (+)-form (Steglich et al., 1974; Osmanova et al., 2010; Gao et al., 2013). Although *T. purpurogenus* is another good producer of diffusible red azaphilone pigments, this species also produce a series of mycotoxins, such as rubratoxin A and B and luteoskyrin in addition to

extrolites that may be toxic if injected intraperitoneally (spiculisporic acid) (Fujimoto et al., 1988) or in the veins of cats (rugulovasine A and B) (Nagaoka et al., 1972; Nagaoka & Kikuchi, 1972). *Talaromyces purpurogenus* can thus not be recommended for industrial production for red pigments.

3.4.1 Taxonomy

Talaromyces atroroseus Yilmaz, Frisvad, Houbraken & Samson *sp. nov.* Figure 3.6. Mycobank MB804901 [urn:lsid:mycobank.org: 804901]

Holotype*:* CBS 133442 in Centraalbureau voor Schimmelcultures is designated as the holotype of *Talaromyces atroroseus*. It was isolated from indoor house dust, Stellenbosch, South Africa by C. Visagie in 2010.

Cultures ex type: CBS 133442 = IBT 32470 = DTO 178A4 = KAS 3778

Etymology: Named after the dark rosy diffusing azaphilone pigment mixture produced.

Diagnosis: Dark green ellipsoidal rough-walled conidia and a dark red diffusing pigment, strains of the species produce the unique combination of secondary metabolites: glauconic acid, ZG–1494α**,** purpuride, red *Monascus* pigments, mitorubrins, and purpactins in fresh isolates.

CYA 25 °C 7d: Colonies are 30–40 mm in diameter, low, plane; margins narrow (1–2 mm), entire, low; mycelia white; texture velvety; sporulation dense, conidia *en masse* dark to dull green; exudate absent; soluble pigment red; reverse colouration dark cherry red.

MEA 25 °C 7d: Colonies 35–40 mm in diameter, low, plane, having a pinkish colour because of exudates diffusing into mycelia; margins narrow (1–2 mm), entire, low; mycelia white; texture velvety overlaying floccose; sporulation moderately dense, conidia *en masse* bluish green; exudate red droplets especially close to margin; soluble pigment absent, after prolonged incubation red pigments produced; reverse colouration dark red.

YES 25 °C 7d: Colonies are 33–45 mm in diameter, raised at centre, sulcate; margins wide (2–3 mm), entire, low; mycelia white; texture velvety; sporulation dense, conidia *en masse* dark to dull green; exudates small red droplets; soluble pigment red in some isolates; reverse colouration brownish red.

CYAS 25 °C 7d: Commonly no growth, some strains up to 5 mm in colony diameter.

CREA 25 °C 7d: Colonies 9–13 mm in diameter, weak acid production close to colony periphery, some strains acid absent; reverse dark red.

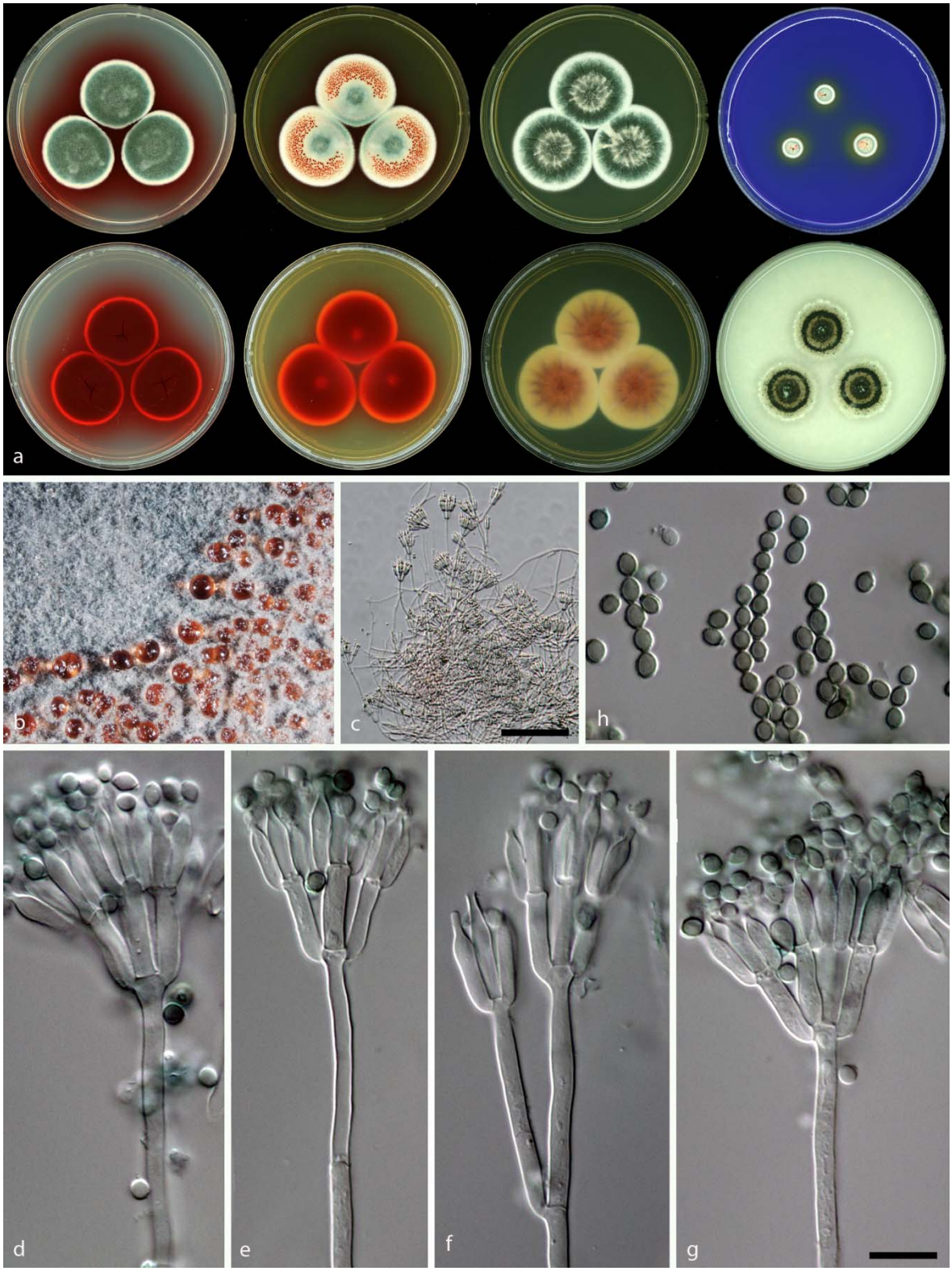


Figure 3.6. Morphological features of *Talaromyces atroroseus* sp. nov. CBS 133442. a: Colonies incubated on CYA, CYA reverse, MEA, MEA reverse, YES, YES reverse, CREA and OA from left to right; b: Colony texture on MEA; c–g: Conidiophores produced on MEA; h: Conidia. (– Scale Bar in c = 50 µm, in g = 10 µm and applies to d–h).

OA 25 °C 7d: Colonies 30–35 mm in diameter, low, plane; margins wide (2–3 mm), entire, low; white mycelia; texture velvety; sporulation dense; conidia *en masse* dull to dark green, almost appears blackish green; exudates absent; soluble pigment absent; reverse colouration commonly greenish yellow to green, red in some isolates.

DG18 25 °C 7d: Colonies 27–30 mm in diameter, low, plane; margins wide (2 mm), entire, low; mycelia white; texture velvety, floccose mycelia present at centre; sporulation dense, conidia *en masse* greyish green, at margins bluish green; exudates absent; soluble pigment absent; reverse colour is beige.

Micromorphology: Conidiophores mostly biverticillate, subterminal branches produced, have a greenish to brownish pigmentation; Stipes smooth walled, 90–150 × 2.5–3 µm; Branches 2–3 when present, 15–50 × 2–3 µm; Metulae in verticils of 3 to 5 per stipe, 8–15 × 3.0–4.0 µm; Phialides acerose, 3 to 6 per metula, 9.5–12.5 × 2.5–3 µm; Conidia rough walled, ellipsoidal, 2–3.5 × 1.5–2.5 µm.

Talaromyces albobiverticillius (H.–M. Hsieh, Y.–M. Ju & S.–Y. Hsieh) Samson, Yilmaz, Frisvad & Seifert, Studies in Mycology 70: 174, 2011. MycoBank MB560683 (Figure 3.7)

Type: BCRC 34774

CYA 25 °C 7d: Colonies 15–20 mm in diameter, low, crateriform, in some isolates sulcate; margins narrow (1–2 mm), entire, low; mycelia white and yellow; texture floccose to velvety; sporulation sparse, in some isolates moderately dense; conidia *en masse* when sparse white, otherwise greyish green; exudates red small droplets; soluble pigmentation red; reverse colouration dark cherry red.

MEA 25 °C 7d: Colonies 24–28 mm in diameter, low, crateriform, in some isolates sulcate; margins wide (2–3 mm), entire, low; mycelia white and yellow; texture velvety with overlaying floccose in the centre; sporulation sparse, in some isolates moderately dense; conidia *en masse* when sparse white, otherwise greyish green; exudates clear and red droplets; soluble red pigment absent; reverse colouration dark red.

YES 25 °C 7d: Colonies 23–25 mm in diameter, raised at centre, sulcate; margins wide (2–3 mm), entire, low; mycelia white and yellow; texture velvety; sporulation sparse, in some isolates moderately dense; conidia *en masse* when sparse white, otherwise greyish green; exudates small orange to red droplets; soluble pigment red in some strains; reverse colouration red to pale brown.

CYAS 25 °C 7d: No growth.

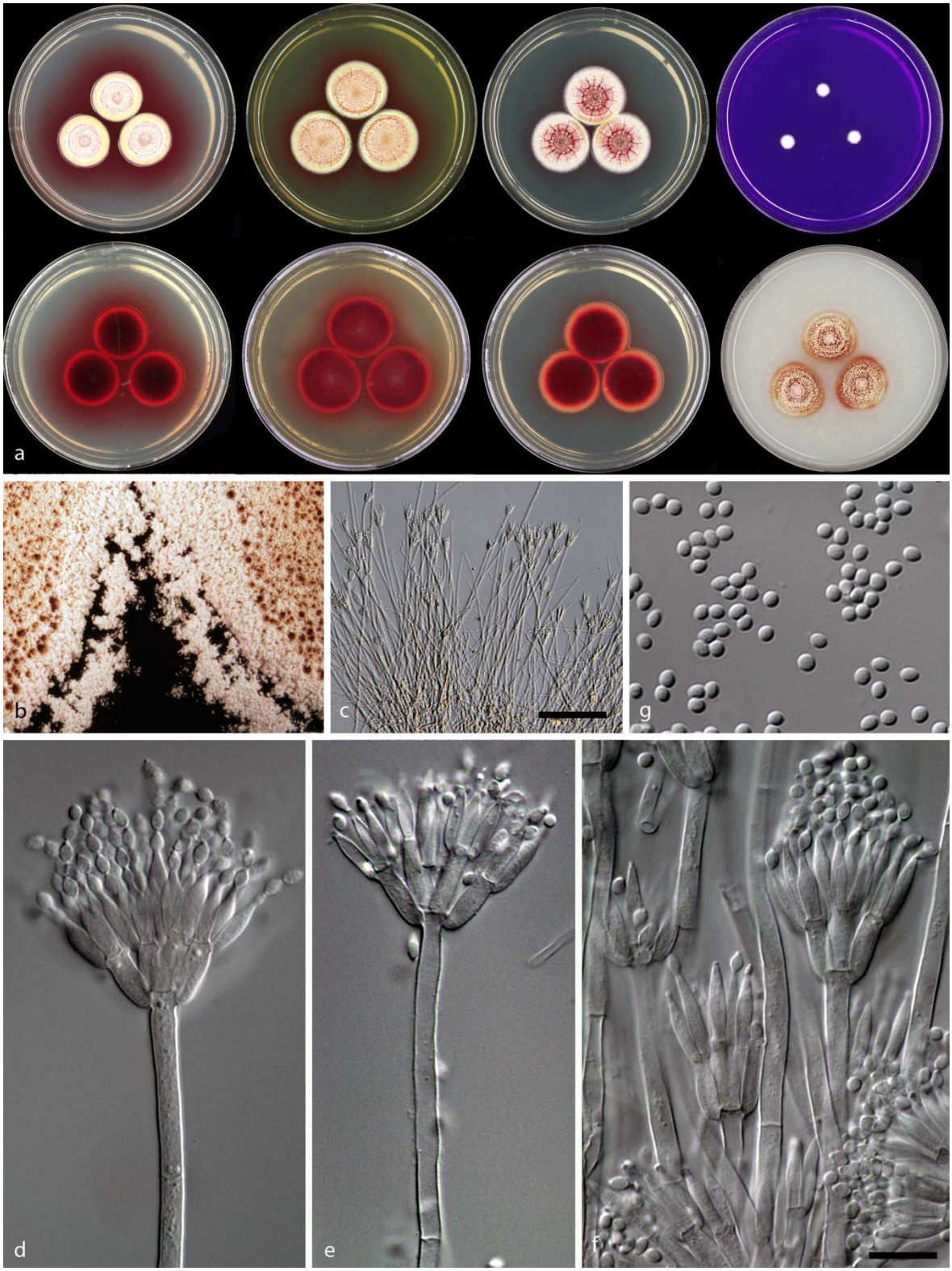


Figure 3.7. Morphological features of *Talaromyces albobiverticillius*. CBS 133440. a: Colonies incubated on CYA, CYA reverse, MEA, MEA reverse, YES, YES reverse, CREA and OA from left to right; b: Colony texture on MEA; c–f: Conidiophores produced on MEA; g: Conidia. (– Scale Bar in c = 50 μm, in f = 10 μm and applies to d–g).

CREA 25 °C 7d: Colonies 4–8 mm in diameter, no acid produced.

OA 25 °C 7d: Colonies 25–28 mm in diameter, low, plane; margins wide (3–4 mm), entire, low; mycelia white; texture velvety; sporulation sparse to moderately dense; conidia *en masse* when sparse white, otherwise greyish green; exudates absent; soluble pigment absent; reverse colouration red in the centre and the rest greenish yellow to green.

DG18 25 °C 7d: Colonies 15–35 mm in diameter, low, plane; margins narrow (1–2 mm), entire, low; mycelia white; texture velvety; sporulation sparse; sparse to moderately dense; conidia *en masse* when sparse white, otherwise greyish green; exudates clear to red droplets; soluble pigment red in some isolates absent; reverse colouration brownish red, in some isolates beige.

Micromorphology: Conidiophores strictly biverticillate, subterminal branches absent; stipes smooth walled, 200– 380 × 2.5–3.5 µm; metulae in verticals of 3–6, 8–12 × 1.5–4.5 µm; phialides acerose, 3–7 per metula, 8–13.5 × 2–3 µm; conidia smooth to finely roughened, spheroid to subglobose, in some isolates fusiform, 2–3.5 (4) × 1.5–2.5 µm.

3.5 Conclusion

Talaromyces atroroseus is a new species that produce large amounts of red pigments that can be potentially used for colouring foods, as it does not produce any known mycotoxins. Certain strains of *T. albobiverticillius* may also be used for these purposes.

Supporting information

Table A.1 is to be found in Appendix A. Table A.1 contains the extrolites searched for by ultra high performance-liquid chromatography-diode array detection-high resolution mass spectrometric detection (UHPLC-DAD-HRMS) in the fungal extracts analysed. The table also includes data on the available standards used in the study.

Acknowledgements

We thank Agilent Technologies for the donation of LC-QTOF and LC-QQQ equipment to DTU Systems Biology via a Thought Leader Award. We also thank Cobus Visagie and Keith A. Seifert for a strain of *Talaromyces atroroseus* (KAS 3778) that served as type of the new species and was isolated from an indoor mold project funded by the Alfred P. Sloan Foundation. Ellen Kirstine Lyhne is thanked for technical assistance.

Author contribution

Conceived and designed the experiments: JCF NY JH RAS. Performed the experiments: JCF NY KBR JH. Analyzed the data: JCF NY UT KBR JH RAS. Contributed reagents/ materials/analysis tools: JCF NY KBR. Wrote the manuscript: JCF NY. Read and improved the draft paper: JCF NY UT KBR JH RAS.

3.6 References

Arai, T., Koganei, K., Umemura, S., Kojima, R., Kato, J., et al., **2013**. Importance of the ammonia assimilation by *Penicillium purpurogenum* in amino derivative *Monascus* pigment, PP-V, production. *AMB Express*, **3**: 19. doi:10.1186/2191-0855-3-19

Arai, T., Umemura, S., Ota, T., Ogihara, J., Kato, J., et al., **2012**. Effects of inorganic nitrogen sources on the production of PP-V [(10Z)-12-carboxyl-monascorubramine] and the expression of the nitrate assimilation gene cluster by *Penicillium* sp. AZ. *Bioscience, Biotechnology, and Biochemistry*, **76**: 120–124. doi:10.1271/bbb.110589

Baldwin, J.E., Barton, D.H., Bloomer, J.L., Jackman, L.M., Rodriguez-Hahn, L., et al., **1962**. The constitutions of glauconic, glaucanic and byssochlamic acids. *Experientia*, **18**: 345–352. doi: 10.1007/BF02172243

Barton, D.H.R., Gondinho, L.D.S. & Sutherland, J.K., **1965**. 331. The nonadrides. Part III. The absolute configuration of glauconic and glaucanic acids. *Journal of the Chemical Society (Resumed)*, **1965**: 1779–1786. doi:10.1039/jr9650001779

Barton, D.H.R., Jackman, L.M., Rodriguez-Hahn, L. & Sutherland, J.K., **1965**. 330. The nonadrides. Part II. The constitutions of glauconic and glaucanic acids. *Journal of the Chemical Society (Resumed)*, **1965**: 1772–1778. doi:10.1039/jr9650001772

Barton, D.H.R. & Sutherland, J.K., **1965**. 329. The nonadrides. Part I. Introduction and general survey. *Journal of the Chemical Society (Resumed)*, **1965**: 1769–1772. doi:10.1039/jr9650001769

Bollinger, P., Haerri, E. & Sigg, H.-P., **1972**. Antibiotic sl 3238 and process for the production of same. US Patent: US3655880 A.

Buechi, G., White, J.D. & Wogan, G.N., **1965**. The structures of mitorubrin and mitorubrinol. *Journal of the American Chemical Society*, **87**: 3484–3489.

Espinoza-Hernández, T.C., Rodriguez-Herrera, R., Aguilar-González, C.N., Lara-Victoriano, F., Reyes-Valdés, M.H., et al., **2013**. Characterization of three novel pigment-producing *Penicillium* strains isolated from the Mexican semi-desert. *African Journal of Biotechnology*, **12**: 3405–3413. doi:10.5897/AJB2013.12338

Feng, Y., Shao, Y. & Chen, F., **2012**. *Monascus* pigments. *Applied Microbiology and Biotechnology*, **96**: 1421–1440. doi:10.1007/s00253-012-4504-3

Frisvad, J.C., Filtenborg, O., Samson, R.A. & Stolk, A.C., **1990**. Chemotaxonomy of the genus *Talaromyces*. *Antonie van Leeuwenhoek International Journal of General and Molecular Microbiology*, **57**: 179–189. doi:10.1007/BF00403953

Frisvad, J.C. & Thrane, U., **1987**. Standardized high-performance liquid chromatography of 182 mycotoxins and other fungal metabolites based on alkylphenone retention indices and UV-VIS spectra (diode array detection). *Journal of Chromatography*, **404**: 195–214. doi:10.1016/0378-8741(88)90090-6

Fujimoto, H., Jisai, Y., Horie, Y. & Yamazaki, M., **1988**. On isolation of spiculisporic acid, a toxic metabolite from *Talaromyces panasenkoi. Proceeding of Japanese Association of Mycotoxicology*, **27**: 15–19.

Gao, J.-M., Yang, S.-X. & Qin, J.-C., **2013**. Azaphilones: chemistry and biology. *Chemical Reviews*, **113**: 4755–4811. doi:10.1021/cr300402y

Gunasekaran, S. & Poorniammal, R., **2008**. Optimization of fermentation conditions for red pigment production from *Penicillium* sp. under submerged cultivation. *African Journal of Biotechnology*, **7**: 1894–1898.

Herrick, H.T. & May, O.E., **1928**. The production of gluconic acid by the *Penicillium luteum-purpurogenum* group. II. Some optimal conditions for acid formation. *Journal of Biological Chemistry*, **77**: 185–195.

Houbraken, J., Due, M., Varga, J., Meijer, M., Frisvad, J.C., et al., **2007**. Polyphasic taxonomy of *Aspergillus* section *Usti. Studies in Mycology*, **59**: 107–128.

Houbraken, J., Lopez-Quintero, C.A., Frisvad, J.C., Boekhout, T., Theelen, B., et al., **2011**. *Penicillium araracuarense* sp nov., *Penicillium elleniae* sp nov., *Penicillium penarojense* sp nov., *Penicillium vanderhammenii* sp nov and *Penicillium wotroi* sp nov., isolated from leaf litter. *International Journal of Systematic and Evolutionary Microbiology*, **61**: 1462–1475.

Houbraken, J. & Samson, R.A., **2011**. Phylogeny of *Penicillium* and the segregation of *Trichocomaceae* into three families. *Studies in Mycology*, **70**: 1–51. doi:10.3114/sim.2011.70.01

Houbraken, J., Spierenburg, H. & Frisvad, J.C., **2012**. *Rasamsonia*, a new genus comprising thermotolerant and thermophilic *Talaromyces* and *Geosmithia* species. *Antonie van Leeuwenhoek International Journal of General and Molecular Microbiology*, **101**: 403–421. doi:10.1007/s10482-011-9647-1

Hsieh, H.-M., Ju, Y.-M. & Hsieh, S.-Y., **2010**. *Penicillium albobiverticillium* sp. nov., a new species producing white conidial masses from biverticillate penicilia. *Fungal Science*, **25**: 25–31.

Huff, R.K., Moppett, C.E. & Sutherland, J.K., **1968**. A novel synthesis of a nine-membered ring. *Chemical Communications (London)*, **1968**: 1192-1193. doi:10.1039/c19680001192

Huff, R.K., Moppett, C.E. & Sutherland, J.K., **1972**. The nonadrides. Part VI. Dimerisation of the C9 unit in vivo and in vitro. *Journal of the Chemical Society, Perkin Transactions 1*, **1972**: 2584–2590. doi:10.1039/p19720002584

Jiang, Y., Li, H.B., Chen, F., Hyde, K.D., **2005**. Production potential of water-soluble *Monascus* red pigment by a newly isolated *Penicillium* sp. *Journal of Agricultural Technology* **1**: 113-126

King, T.J., Roberts, J.C. & Thompson, D.J., **1970**. Structure of purpurogenone, a metabolite of *Penicillium purpurogenum* Stoll: an X-ray study. *Journal of the Chemical Society D-Chemical Communications*, **1970**: 1499. doi:10.1039/C2970001499A

King, T.J., Roberts, J.C. & Thompson, D.J., **1973**. Studies in mycological chemistry. Part XXX and last. Isolation and structure of purpuride, a metabolite of *Penicillium purpurogenum* Stoll. *Journal of the Chemical Society. Perkin Transactions 1*, **1**: 78–80. doi:10.1039/p19730000078

Kondo, H., Kurama, M., Nakajima, S., Osada, K. & Okura, A., **1993**. New substance BE-25327 which is an estradiol agonist – useful for treating oestrogen-deficient gynaecological diseases, osteoporous, prostatic cancer and prostatomegaly. Japanese patent: JP5032579-A.

Lin, Y.-L., Wang, T.-H., Lee, M.-H. & Su, N.-W., **2008**. Biologically active components and nutraceuticals in the *Monascus*-fermented rice: a review. *Applied Microbiology and Biotechnology*, **77**: 965–973. doi:10.1007/s00253-007-1256-6

Liu, B.-H.H., Wu, T.-S.S., Su, M.-C.C., Chung, C.P. & Yu, F.-Y.Y., **2005**. Evaluation of citrinin occurrence and cytotoxicity in *Monascus* fermentation products. *Journal of Agricultural and Food Chemistry*, **53**: 170–175. doi:10.1021/jf048878n

Mapari, S.A.S., Hansen, M.E., Meyer, A.S. & Thrane, U., **2008**. Computerized screening for novel producers of *Monascus*-like food pigments in *Penicillium* species. *Journal of Agricultural and Food Chemistry*, **56**: 9981–9989. doi:10.1021/jf801817q

Mapari, S.A.S., Meyer, A.S., Thrane, U. & Frisvad, J.C., **2009**. Identification of potentially safe promising fungal cell factories for the production of polyketide natural food colorants using chemotaxonomic rationale. *Microbial Cell Factories*, **8**: 24. doi:10.1186/1475-2859-8-24

Mapari, S.A.S., Nielsen, K.F., Larsen, T.O., Frisvad, J.C., Meyer, A.S., et al., **2005**. Exploring fungal biodiversity for the production of water-soluble pigments as potential natural food colorants. *Current Opinion in Biotechnology*, **16**: 231–238. doi:10.1016/j.copbio.2005.03.004

Mapari, S.A.S., Thrane, U. & Meyer, A.S., **2010**. Fungal polyketide azaphilone pigments as future natural food colorants? *Trends in Biotechnology*, **28**: 300–307. doi:10.1016/j.tibtech.2010.03.004

Marsini, M.A., Gowin, K.M. & Pettus, T.R.R., **2006**. Total synthesis of (+/-)-mitorubrinic acid. *Organic Letters*, **8**: 3481–3483.

Méndez, A., Pérez, C., Montañéz, J.C., Martínez, G., Aguilar, C.N.C.N., et al., **2011**. Red pigment production by *Penicillium purpurogenum* GH2 is influenced by pH and temperature. *Journal of Zhejiang University. Science. B*, **12**: 961–968. doi:10.1631/jzus.B1100039

Mendez-Zavala, A., Contreras-Esquivel, J.C., Lara-Victoriano, F., Rodriguez-Herrera, R. & Aguilar, C.N., **2007**. Fungal production of the red pigment using a xerophilic strain *Penicillium purpurogenum* GH-2. *Revista Mexicana De Ingenieria Quimica*, **6**: 267–273.

Nagaoka, A. & Kikuchi, K., **1972**. Pharmacological studies of new indole alkaloids, rugulovasine A and B hydrochloride. II. Hypotensive mechanism of both alkaloids in anesthetized cats. *Arzneimittel-Forschung*, **22**: 143–146.

Nagaoka, A., Nagawa, Y. & Kikuchi, K., **1972**. Pharmacological studies of new indole alkaloids, rugulovasine A and B hydrochloride. I. Effects of both alkaloids on cardiovascular and central nervous-system and smooth muscles. *Arzneimittel-Forschung*, **22**: 137–142.

Natsume, M., Takahashi, Y. & Marumo, S., **1985**. (-)-Mitorubrinic acid, a morphogenic Substance inducing chlamydospore-like cells, and its related new metabolite, (+)-mitorubrinic acid-B, isolated from *Penicillium funiculosum*. *Agricultural and Biological Chemistry*, **49**: 2517–2519.

Nielsen, K.F.K., Månsson, M., Rank, C., Frisvad, J.C., Larsen, T.O., et al., **2011**. Dereplication of microbial natural products by LC-DAD-TOFMS. *Journal of natural products*, **74**: 2338–2348. doi:10.1021/np200254t

Nishida, H., Tomoda, H., Cao, J., Okuda, S. & Omura, S., **1991**. Purpactins, new inhibitors of acyl-CoA - cholesterol acyltransferase produced by *Penicillium-purpurogenum*. II. Structure elucidation of purpactin-A, purpactin-B and purpactin-C. *Journal of Antibiotics*, **44**: 144–157.

Ogihara, J., Kato, J., Oishi, K. & Fujimoto, Y., **2001**. PP-R, 7-(2-hydroxyethyl)-monascorubramine, a red pigment produced in the mycelia of *Penicillium* sp. AZ. *Journal of Bioscience and Bioengineering*, **91**: 44–47.

Ogihara, J., Kato, J., Oishi, K., Fujimoto, Y. & Eguchi, T., **2000**. Production and structural analysis of PP-V, a homologue of monascorubramine, produced by a new isolate of *Penicillium* sp. *Journal of Bioscience and Bioengineering*, **90**: 549–554.

Ogihara, J. & Oishi, K., **2002**. Effect of ammonium nitrate on the production of PP-V and monascorubrin homologues by *Penicillium* sp. AZ. *Journal of Bioscience and Bioengineering*, **93**: 54–59.

Osmanova, N., Schultze, W. & Ayoub, N., **2010**. Azaphilones: a class of fungal metabolites with diverse biological activities. *Phytochemistry Reviews*, **9**: 315–342. doi:10.1007/s11101-010-9171-3

Patakova, P., **2013**. Monascus secondary metabolites: Production and biological activity. *Journal of Industrial Microbiology and Biotechnology*, **40**: 169–181. doi:10.1007/s10295-012-1216-8

Pharm Amano, **1974**. Novel pigment Red W59 extracted from *Penicillium purpurogenum* culture medium with organic solvents. Japanese patent: JP 49093587-A.

Pitt, J.I., **1980**. *The genus Penicillium and its teleomorphic states Eupenicillium and Talaromyces.*, New York: Academic Press. p634

Qiu, M., Xie, R., Shi, Y., Chen, H., Wen, Y., et al., **2010**. Isolation and identification of endophytic fungus SX01, a red pigment producer from *Ginkgo biloba* L. *World Journal of Microbiology and Biotechnology*, **26**: 993–998. doi:10.1007/s11274-009-0261-6

Raper, K.B. & Thom, C., **1949**. *A manual of the penicillia, by Kenneth B. Raper and Charles Thom; with the technical assistance and illus. by Dorothy I. Fennel.*, Baltimore: Williams & Wilkins Co. 875 pp. doi:10.5962/bhl.title.4993

van Reenen-Hoekstra, E.S., Frisvad, J.C., Samson, R.A. & Stolk, A.C., **1990**. The *Penicillium funiculosum* complex - well defined species and problematic taxa. In R. A. Samson & J. I. Pitt, eds. *Modern Concepts in Penicillium and Aspergillus Classification*. pp. 173–192. doi:10.1007/978-1-4899-3579-3_15

Roberts, J.C. & Thompson, D.J., **1971a**. Studies in mycological chemistry. Part XXVIII. Isolation and structure of deoxypurpurogenone, a minor pigment of *Penicillium purpurogenum* Stoll. *Journal of the Chemical Society. Perkin Transactions 1*, **20**: 3493–3495. doi:10.1039/j39710003493

Roberts, J.C. & Thompson, D.J., **1971b**. Studies in myocological chemistry. XXVII. Reinvestigation of the structure of purpurogenone, a metabolite of *Penicillium purpurogenum* Stoll. *Journal of the Chemical Society. Perkin Transactions 1*, **1971**: 3488–3492. doi:10.1039/j39710003488

Roberts, J.C. & Warren, C.W.H., **1955**. Studies in mycological chemistry. Part IV. Purpurogenone, a metabolic product of *Penicillium purpurogenum* Stoll. *Journal of the Chemical Society (Resumed)*, **1955**: 2992. doi:10.1039/jr9550002992

Samson, R.A., Houbraken, J., Thrane, U., Frisvad, J.C. & Andersen, B., **2010**. *Food and indoor fungi*, CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands.

Samson, R.A., Stolk, A.C. & Frisvad, J.C., **1989**. Two new synnematous species of *Penicillium*. *Studies in Mycology*, **31**: 133–143.

Samson, R.A., Yilmaz, N., Houbraken, J., Spierenburg, H., Seifert, K. a, et al., **2011**. Phylogeny and nomenclature of the genus *Talaromyces* and taxa accommodated in *Penicillium* subgenus *Biverticillium*. *Studies in Mycology*, **70**: 159–183. doi:10.3114/sim.2011.70.04

Santos-Ebinuma, V.C., Roberto, I.C., Teixeira, M.F.S. & Jr., A.P., **2013a**. Improving of red colorants production by a new *Penicillium purpurogenum* strain in submerged culture and the effect of different parameters in their stability. *Biotechnology Progress*, **29**: 778–785.

Santos-Ebinuma, V.C., Teixeira, M.F.S. & Jr., A.P., **2013b**. Submerged culture conditions for the production of alternative natural colorants by a new isolated *Penicillium purpurogenum* DPUA 1275. *Journal of Microbiology and Biotechnology*, **23**: 802–810.

de Silva, E.D., Williams, D.E., Jayanetti, D.R., Centko, R.M., Patrick, B.O. et al., **2011.** Dhilirolides A-D, meroterpenoids produced in culture by the fruit-infecting fungus *Penicillium purpurogenum* collected in Sri Lanka. *Organic Letters* **13**: 1174-1177. doi:10.1021/ol200031t.

Steglich, W., Klaar, M. & Furtner, W., **1974**. (+)-Mitorubrin derivatives from *Hypoxylon fragiforme*. *Phytochemistry*, **13**: 2874–2875. doi:10.1016/0031-9422(74)80262-1

Stoll, O., **1903-1904**. *Beiträge zur morphologischen und biologischen Charakteristik von Penicilliumarten*. Dissertation. Würzburg.

Suzuki, S., Hosoe, T., Nozawa, K., Yaguchi, T., Udagawa, S., et al., **1999**. Mitorubrin derivatives on ascomata of some *Talaromyces* species of ascomycetous fungi. *Journal of natural products*, **62**: 1328–1329. doi:10.1021/np990146f

Takashima, M., Kitajama, A. & Otuka, K., **1955**. Studies on the pigment of a white crystal "glauconic acid" and red pigments by *Penicillium purpurogenum*. *Journal of the Agricultutal Chemical Society of Japan*, **29**: 25–29.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., et al., **2011**. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**: 2731–2739. doi:10.1093/molbev/msr121

Tomoda, H., Nishida, H., Masuma, R., Cao, J., Okuda, S., et al., **1991**. Purpactins, new inhibitors of acyl-CoA - cholesterol acyltransferase produced by *Penicillium purpurogenum*. I. Production, isolation and physicochemical and biological properties. *Journal of Antibiotics*, **44**: 136–143.

Velmurugan, P., Kamala-Kannan, S., Balachandar, V., Lakshmanaperumalsamy, P., Chae, J.-C., et al., **2010a**. Natural pigment extraction from five filamentous fungi for industrial applications and dyeing of leather. *Carbohydrate Polymers*, **79**: 262–268. doi:10.1016/j.carbpol.2009.07.058

Velmurugan, P., Lee, Y.H., Venil, C.K., Lakshmanaperumalsamy, P., Chae, J.-C., et al., **2010b**. Effect of light on growth, intracellular and extracellular pigment production by five pigment producing filamentous fungi in synthetic medium. *Journal of Bioscience and Bioengineering*, **109**: 346–350. doi:10.1016/j.jbiosc.2009.10.003

Wang, H., Wang, Y., Wang, W., Fu, P., Liu, P., et al., **2011**. Anti-influenza virus polyketides from the acid-tolerant fungus *Penicillium purpurogenum* JS03-21. *Journal of Natural Products*, **74**: 2014–2018. doi:10.1021/np2004769

West, R.R., VanNess, J., Varming, A.M., Rassing, B., Biggs, S., et al., **1996**. ZG-1494 alpha, a novel platelet-activating factor acetyltransferase inhibitor from *Penicilium rubrum*, isolation, structure elucidation and biological activity. *Journal of Antibiotics*, **49**: 967–973.

Wijkman, N., **1931**. On some new substances made through mould-fungus. *Justus Liebigs Annalen der Chemie*, **485**: 61–73.

Yilmaz, N., Houbraken, J., Hoekstra, E.S., Frisvad, J.C., Visagie, C.M., et al., **2012**. Delimitation and characterisation of *Talaromyces purpurogenus* and related species. *Persoonia*, **29**: 39–54. doi:10.3767/003158512X659500

Yuill, J.L., **1934**. The acids produced from sugar by a *Penicillium* parasitic upon *Aspergillus niger*. *Biochemical Journal*, **28**: 222–227.

Chapter 4

Pigment profile of *Talaromyces atroroseus* IBT11181 from submerged fermentation

Part of the overall project was to further investigate the potential of *Talaromyces* as production strain for red *Monascus* pigments. This involves transferring the cultivation setup from shake flasks to bioreactors for further optimization of cultivation parameters.

In the present study initial investigation of the potential for reactor-based submerged-fermentation of *Talaromyces atroroseus* as production organism for red *Monascus* pigments is presented. The aims was to chemical profile the produced *Monascus* pigments during submerged fermentation as well as investigate the dynamics of the production of the individual pigments, i.e. whether different pigments are produced at different stages of the fermentation process. Further optimization of fermentation setup as well as morphological characterisation is carried out by PhD student Gerit Nymschefsky and will hence not be addressed in this thesis.

During this study one novel *Monascus* pigment was purified and structure elucidated. A preliminary manuscript of the purification and elucidation of the novel compound can be found in appendix B.

4.1 Introduction

Talaromyces atroroseus and related *Talaromyces* species are promising candidates for large scale mycotoxin free production of water-soluble red *Monascus* pigments (MPs) (Frisvad et al., 2013). Several steps towards physiological characterization and optimization of pigment yields have already been taken by others.

In several studies the red pigment production by *Pencillium purpurogenum* GH2 (Mendez-Zavala et al., 2007; Méndez et al., 2011; Espinoza-Hernández et al., 2013) and *P. purpurogeunm* DPUA1275 (Santos-Ebinuma et al., 2013a; Santos-Ebinuma et al., 2013b) is described. Both *P. purpurogenum* GH2 and *P. purpurogenum* DPUA 1275 are potential *T. atroroseus* strains according to the taxonomic revision of *Penicillium* subgenus *Biverticillium* (Samson et al., 2011) and the taxonomic description of *T. atroroseus* (Chapter 3 and Frisvad et al., 2013). The studies of the GH2 and DPUA 1275 strains has however only been carried out as shake flask cultivations, and furthermore no growth rate has been calculated and no metabolite profiling have been performed.

In shake flask the control of the process is very limited. The aim of the present study was to investigate the feasibility of submerged fermentation in bioreactors under controlled conditions. Furthermore and most important for my part of the study was to investigate the pigment production profiles produced in the batch fermentation.

4.2 Materials and methods

4.2.1 Strains and solid media

All fungal strains used in this study were obtained from the IBT culture collection at Department of Systems Biology, Technical University of Denmark. For screening of red pigment production on solid media standard mycological media were prepared according to Samson *et al.* (Samson et al., 2010).

4.2.2 Media for liquid cultivations

Two different media was used for liquid cultivations. For shake flask cultivations the complex CY medium was prepared as the solid CYA medium except for the agar. Specifically the final medium composition was: 35 $g.L^{-1}$ Czapek Dox Broth (Difco, giving final

concentrations of: 30 g.L^{-1} sucrose; 3 g.L^{-1} $NaNO_3$; 1 g.L^{-1} K_2PO_4; 0.5 g.L^{-1} KCl; 0.5 g.L^{-1} $MgSO_4$; 0.01 g.L^{-1} $FeSO_4$); 5 g.L^{-1} Yeast extract; 0.01 g.L^{-1} $ZnSO_4 \cdot 7H_20$; 0.005 g.L^{-1} $CuSO_4 \cdot 5H_2O$). pH was adjusted to 6.2 prior to autoclavation.

For reactor fermentations a defined medium with the following composition was used (Nymschefsky, unpublished): 20 g.L^{-1} sucrose; 1 g.L^{-1} glucose; 10 g.L^{-1} KNO_3; 8 g.L^{-1} K_2PO_4; 1 g.L^{-1} NaCl; 2 g.L^{-1} $MgSO_4$; 0.5 g.L^{-1} KCl; 0.1 g.L^{-1} $CaCl \cdot 2H_20$; 1 mL.L^{-1} trace metal solution (composed of: 0.4 g.L^{-1} $CuSO_4 \cdot 5H_2O$; 0.04 g.L^{-1} $Na_2B_4O_7 \cdot 10H_2O$; 0.8 g.L^{-1} $FeSO_4 \cdot 7H_2O$; 0.8 g.L^{-1} $MnSO_4 \cdot 2H_2O$; 0.8 g.L^{-1} $Na_2MoO_4 \cdot 2H_2O$; and 0.8 g.L^{-1} $ZnSO_4 \cdot 7H_2O$).

4.2.3 Inoculum preparation

Cultures were propagated on CYA medium for 7 days and subsequently harvested with 0.9 % NaCl and filtered through sterile mira cloth. Spore concentration in the spore suspensions were evaluated by hemocytometer and cultivations initiated by inoculation of the spore suspension to a final concentration of $1 \cdot 10^6$ spores.L^{-1}.

4.2.4 Shake flask cultivations

For shake flask cultivations the complex CY medium. Cultures were inoculated in 500 mL unbaffled shake flasks containing 100 mL of the CY medium. Cultivations were run for seven days at 30 °C with shaking at 150 rpm. All shake flasks were run in duplicates.

4.2.5 Batch cultivations in bioreactors

Batch cultivations were carried out in 2 L Sartorius Biostat B bioreactors with a working volume of 1.8 L. The bioreactors were equipped with two Rushton six-blade disc turbines for proper agitation. The pH was monitored with a calibrated pH electrode (Mettler Toledo, Denmark).

The temperature was kept at 30 °C throughout the cultivation and pH maintained at pH 5.0 by automatic addition of 2 M NaOH and 2 M H_2SO_4. The bioreactor was sparged with sterile atmospheric air. Batch fermentation was run in duplicates.

4.2.6 Sampling and biomass determination

Samples for biomass determination as well as for pigment profiling and quantification was taken from the bioreactor cultivations starting at the beginning of the exponential phase and until termination of the fermentation.

For biomass determination ~4-6 mL sample was quickly withdrawn from the culture. The exact sample size was determined by weight before the sample was filtered through a pre-dried 0.45 μm polyether sulfone filter (Frisenette, Knebel, Denmark) and washed thoroughly with water. The filter containing the biomass was dried in microwave oven at 150 W for 20 min and cooled in a desiccator. Dry weight (DW) concentrations of the biomass were subsequently determined gravimetrically.

Samples for pigment profiling and quantification were filtered through a 0.45 μm cellulose acetate filter (Frisenette; Knebel; Denmark) and flash frozen in liquid nitrogen. Samples were wrapped in aluminium foil to shield for light-degradation and stored at -20 °C until further analysis.

4.2.7 Quantification of carbon sources

Concentration of sugars, i.e. sucrose, glucose and fructose was performed on a Dionex Ultimate 3000 ultra high performance liquid chromatography (UHPLC) system equipped with a Shodex RI-101 refractive index detector. Separation of the sugars was done on an Aminex HPX-87H Ion Exclusion column with the dimension 300mm x 7.8mm (BioRad, Hercules, Ca, USA). The chromatography was run in isocratic mode with 5 mM H_2SO_4 as mobile phase. The separation was run at 30 °C at a flow rate of 0.5 $mL.min^{-1}$. Concentration of sucrose, glucose and fructose were calculated using six-point external standard curves with concentrations ranging between 1-20 $g.L^{-1}$.

4.2.8 Quantification of overall red pigments

For quantification of the overall red pigment production absorbance spectra were measured with a Synergy 2 Multi-Mode plate reader (BioTek, Winooski, Vermont, USA) reading from 200 nm to 700 nm with 5 nm steps. Specifically 100 μl filtered samples from the batch fermentation were measured in a 96 Well Microplate with a standard flat bottom (F/ST) well design (Greiner Bio-One, Frickenhausen, Germany). The resulting path length measured was 2.90 mm.

4.2.9 Chromatographic analysis

Filtered samples were analysed by UHPLC coupled to high-resolution mass spectrometry (HRMS). Liquid chromatography was performed on an Agilent 1290 Infinity LC system with a diode array detector (DAD) coupled to an Agilent 6550 iFunnel Q-TOF with an electrospray

ionization source. The separation was performed on a 2.1 x 250 mm, 2.7 μm Poroshell 120 Phenyl-Hexyl column (Agilent) at 60 °C with a water-acetonitrile gradient (both with 20 mM formic acid) going from 10 % (v/v) to 100 % acetonitrile in 15 min followed by 2.5 min with 100 % acetonitrile and then returning to the start conditions for 2.5 min for equilibration before next sample. Injection volume for all samples was 4 μl.

4.2.10 Data processing

The software program Qualitative Analysis B.06.00 from the Agilent Masshunter software suite was used for processing of the acquired analytical data. Pigmented compounds were identified in extracted wave length chromatograms (EWC). The wavelengths used correspond to absorption maxima of the different MP chromophore classes, i.e. 390±4 nm, 470±4 nm and 525±4 nm. Dereplication of individual compounds were based on assigned $[M+H]^+$ ion masses and confirmed by the support of $[M+Na]^+$ ions as well as UV/VIS spectra.

Due to closely eluting peaks the EWCs did not reach the baseline between all peaks. Hence peak heights were used for relative quantitation of individual pigments. To get reliable peak heights the Chemstation Integrator were optimized so the integration baseline fitted with the baseline of the background sample, i.e. a sample taken directly after inoculation of the reactors. The resulting height of the integrated area for each peak was used as a relative quantitative measurement of the individual pigments. Since no authentic standards for most of the produced compounds were available the quantity of each compound is expressed as peak height under the given chromatographic conditions.

4.3 Results and discussion

4.3.1 Initial screening of pigment production

A selection of *T. atroroseus* strains as well as two other *Talaromyces* strains was initially screened for pigment production when grown for 7 days at 30 °C on a range of common fungal cultivation media (See Figure 4.1).

Talaromyces sp. IBT14065 did not produce any red pigmentation on any of the investigated media and was hence discarded as suitable organism for pigment production. *Talaromyces australis* IBT14256 showed slightly red colouration on the reverse on especially CYA and

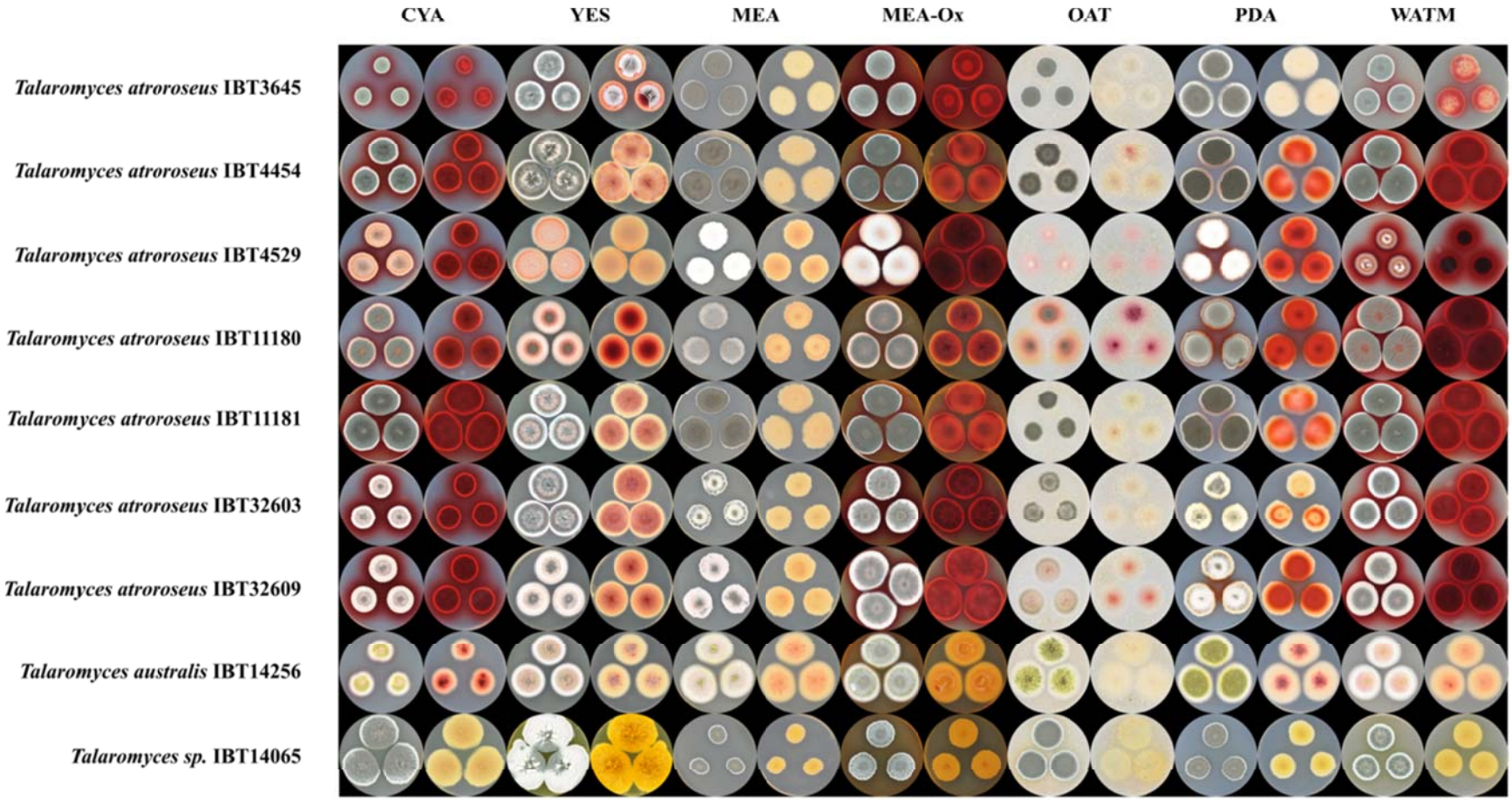


Figure 4.1. Solid media screening of *Talaromyces* strains for production of red *Monascus* pigments.
Talaromyces sp. IBT14065 has previously been taxonomic classified as *Penicillium monascum* (Mapari et al., 2009) and *Penicillium funiculosum* (Mapari et al., 2010) however following the movement of *Penicillium* subgenus biverticillium into *Talaromyces* (Samson et al., 2011) along with ongoing taxonomic revisions of the genus *Talaromyces* (Yilmaz, 2015) this strains needs further examination to potentially identify it as new species (Frisvad, personal communication).

PDA, however not nearly in the same range as the *T. atroroseus* strains. Prolonged incubation of up to 14 days of *T. australis* IBT14256 however significantly increased the production of red pigments (data not shown). However the long incubation time before pigment production occurred together with less prior knowledge of potential production of mycotoxins in *T. australis* made me discard the strain for further investigations.

It is clear that MP production in *T. atroroseus* strain is highly dependent by the growth media. Especially the red pigments production capability on CYA, MEA-Ox and WATM are excellent, while no production of red pigments on MEA is observed in any of the strains. The strains IBT11180 and IBT4529 are two different culture numbers which both can be traced back to strain FRR1061 numbered by J. I. Pitt. Original this strain was isolated by K. Saito in 1918 in Manchuria, China. The FRR1061 strain has been received and deposited to the IBT collection two times, giving rise to two individual strain numbers. Comparing the morphological appearance of these two strains it is evident that there are remarkable differences. IBT4529 show very poor sporulation on all media; a trait that has often been observed in old cultures of *Talaromyces* (Yilmaz, 2015).

Among the other *T. atroroseus* strains IBT3645 seems to be less suitable for red pigment production especially on CYA and WATM where it produces significantly less pigments than the other strains. This could be linked to the smaller colony diameter also observed on the same media, identifying it as a strain with slower growth rate.

Among the other strains the observed pigment production on plates is not usable for distinguishing between the potential for pigment production. Hence shake flask experiments investigating the potential for pigment production in liquid cultures were set up. CYA is a good solid medium for pigment production and CY liquid medium has in a previous study shown promising results for pigment production in liquid cultivations using LECA as solid support (Mapari et al., 2009b). Based on these observations screening of pigment production potential of different *T. atroroseus* strains were evaluated in shake flask with CY as liquid medium. The cultivations were run for seven days, but for all strains only very faint reddish hue was observed.

The pigment production potential under the given conditions was not very promising. Hence it was decided to choose one *T. atroroseus* strain to use for optimization of culture and media conditions. The choice fell on *T. atroroseus* IBT11181 since it in previous studies at our Department had shown some promising results in relation to pigment production (Mapari, 2008; Mapari et al., 2009a; Mapari et al., 2009b) and at the same time was in the process of

being genome sequenced (Chapter 5) Media and culture condition optimization was done by PhD student Gerit Nymschefsky and the evaluation and results of these will not be discussed here. The final media for liquid cultivation share similarities with the CY medium since it is based on sucrose and KNO_3 as carbon- and nitrogen-source respectively just as CY. Before the reactor cultivations azaphilone pigment profiling by *T. atroroseus* IBT11181 grown on solid CYA was performed.

4.3.2 Azaphilone production by *T. atroroseus* IBT11181 on CYA

The genus of *Talaromyces* is well-known for the production of a series of yellow azaphilones (Samson et al., 2011; Frisvad et al., 1990) which has a O-orsellinic acid or O-orsellinic acid derivative attached to the bicyclic azaphilone core. These compounds collectively known as mitorubrins include among others mitorubrin, mitorubrinol, mitorubrinal, mitorubrinic acid, purpurquinone A-C (Frisvad et al., 1990; Samson et al., 2011; Wang et al., 2011). Many of the *Talaromyces* species also produce the red azaphilone pigments referred to as *Monascus* pigments (MP) and among these species is *Talaromyces atroroseus* (Chapter 3 and Frisvad et al., 2013). *Talaromyces atroroseus* IBT11181 is a particular good azaphilone producer with a high degree of diffusible red pigments. Previously *Talaromyces atroroseus* IBT11181 has been reported to produce a number of azaphilones including the MPs N-glutarylmonascorubramine, N-glutarylrubropunctamine, monascorubrin, PP-O and PP-R (Mapari et al., 2009b; Chapter 3 and Frisvad et al., 2013), as well as mitorubrin, mitorubrinol and purpurquinone A (Chapter 3 and Frisvad et al., 2013).

To compare the pigment production between solid CYA medium and the defined liquid medium the azaphilone pigments produced by *T. atroroseus* IBT11181 on solid CYA were identified. The identification of produced azaphilones was performed using the UV/VIS spectral fingerprint for the different expected azaphilone classes as described in Chapter 2. More specifically I identified azaphilone chromophores as peaks in the extracted wavelength chromatograms (EWC) at 350 nm for the mitorubrins and yellow MPs, at 490 nm for the orange MPs and at 525 nm for the red MPs, see Figure 4.3. For compounds with UV/VIS spectra reflecting known azaphilone chromophores I dereplicated the masses and assigned products when possible. In the following bold numbers will refer to the numbering in Figure 4.3-4.4.

With the described approach I were not able to identify any compounds with the classical mitorubrin chromophore however I identified Purpurquinone A (**3**), a dihydroxylated

derivative of mitorubrin (Wang et al., 2011). Compared to mitorubrin the dihydroxylation in Purpurquinone A has removed one of the conjugated double bounds giving rise to the different UV/VIS spectrum. The observed spectrum is consistent with the Purpurquinone A spectrum described by Wang *et al.* (Wang et al., 2011). I also identified two small azaphilone compounds, **1** and **2.** The UV/VIS spectrum of **2** is very similar to monascin and ankaflavin thus must contain the same structural chromophore. With $[M+H]^+$=237.1126 the mass of **2** match FK17-P2b1 and FK17-P2b2 which was firstly identified in *Aspergillus sp.* (Arai & Sano 1994). FK17-P2b1 has since been identified as an intermediate in the *Monascus* pigment biosynthesis in *Monascus kaoliang* and *Monasus purpureus* (Jongrungruangchok et al., 2004; Bijinu et al., 2014), although Jongrungruangchok *et al.* wrongly refers to the compound as FK17-P2b2. The structural chromophore of FK17-P2b1 is identical to that of monascin and ankaflavin (Figure 4.2) so based on the UV/VIS spectrum compound **2** is believed to be FK17-P2b1.

Compound **1** with the mass ion $[M+H]^+$=255.1234 match the compound Monascusone A (Figure 4.4), which is similar to FK17-P2b1 but with the double bond towards the end of the polyketide hydrated. Monascusone A was originally identified as an intermediate in MP biosynthesis in the same study as FK17-P2b1 (Jongrungruangchok et al., 2004).

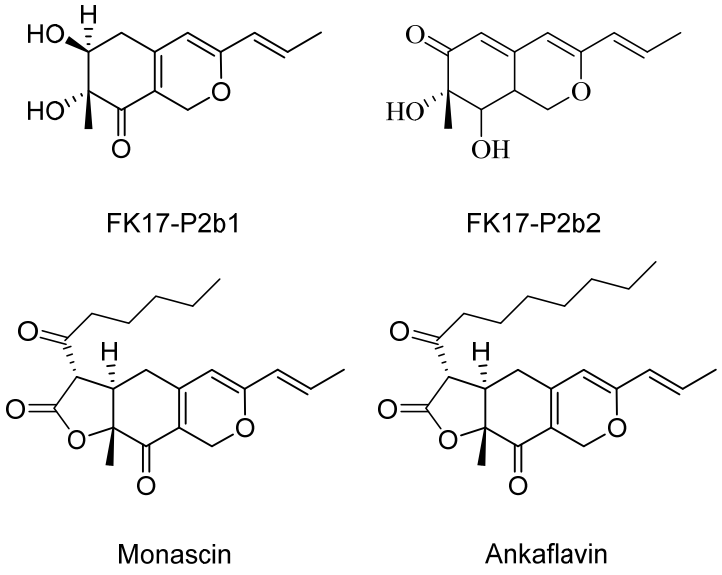


Figure 4.2. Structures of FK17-P2b1, FK17-P2b2, monascin and ankaflavin. FK17-P2b1 exhibits the same chromophore as monascin and ankaflavin indicating that the UV/VIS spectrum should be similar.

Using the extracted wavelength chromatogram at 490 nm (not shown in Figure 4.3) I verified the production of the orange *Monascus* pigment monascorubrin (**8**). Looking into the compounds responsible for the red pigmentation I found many peaks having the characteristic nitrogen-containing MP double UV/VIS peaks around 430 and 525 nm. Based on this unique fingerprint all of these compounds have the same chromophore as monascorubramine and rubropunctamine as described in Chapter 2. Several of these peaks were however very small and represents minor compounds. Therefore it was not possible to assign masses and chemical formulas to most of these red compounds.

As all the red compounds have the characteristic UV/VIS profile of the nitrogen-containing MP chromophore I used the criterion of minimum one nitrogen in the compound in the assignment of possible chemical formulas. For the compounds present in higher quantities I positively identified N-threonyl rubropunctamine (**4**) (also described as the threonine derivative of rubropunctatin) as well as PP-R (**6**) (Ogihara et al., 2001).

Apart from the known compounds I also linked three red unknown *Monascus* pigments (**5,7,9**) to their masses. For compound **5** showing the mass ion $[M+H]^+$=532.1984 supported by the sodium adduct $[M+Na]^+$=554.1798 I propose the chemical formula to be $C_{30}H_{29}NO_8$ with a mass difference between the observed and calculated mass of -3.4 ppm referring to the $[M+H]^+$ ion. For compound **7** showing the mass ion $[M+H]^+$=760.3709 supported by the sodium adduct $[M+Na]^+$=782.3527 I propose the chemical formula to be $C_{46}H_{51}N_2O_{11}$ with a mass difference between the observed and calculated mass of 1.21 ppm referring to the $[M+H]^+$ ion. To my knowledge no *Monascus* pigment with the proposed chemical formulas of any of these two compounds has previously been described. I speculate that compound **7** could be a dimeric derivative of monascorubramin. Dimers of azaphilones has previously been described for mitorubrins in the form of (-)-diazaphilonic acid (Tabata et al., 1999) and the chaetoglobin A and B, produced by a *Chaetomium* strain (Ming Ge et al., 2008). The dimer of monascorubramine, with the same 5-5' linkage as seen in the chaetoglobins, results in a molecular formula of $C_{46}H_{52}N_2O_8$ which differs from the proposed formula by having one extra hydrogen and three fewer oxygen. The observed differences could be accounted for by oxidations at various sites, however with the data available it is not possible to further resolve this. Compound **9** was firstly dereplicated via $[M_1+H]^+$=784.5864 supported by the sodium adduct $[M_1+Na]^+$=806.5685 to have the formula $C_{51}H_{77}NO_5$. However a closer look of the extracted ion chromatogram (EIC) for the mentioned masses (see appendix E) revealed that the shape and time of the EIC peak was slightly different from the EWC_{525nm} peak.

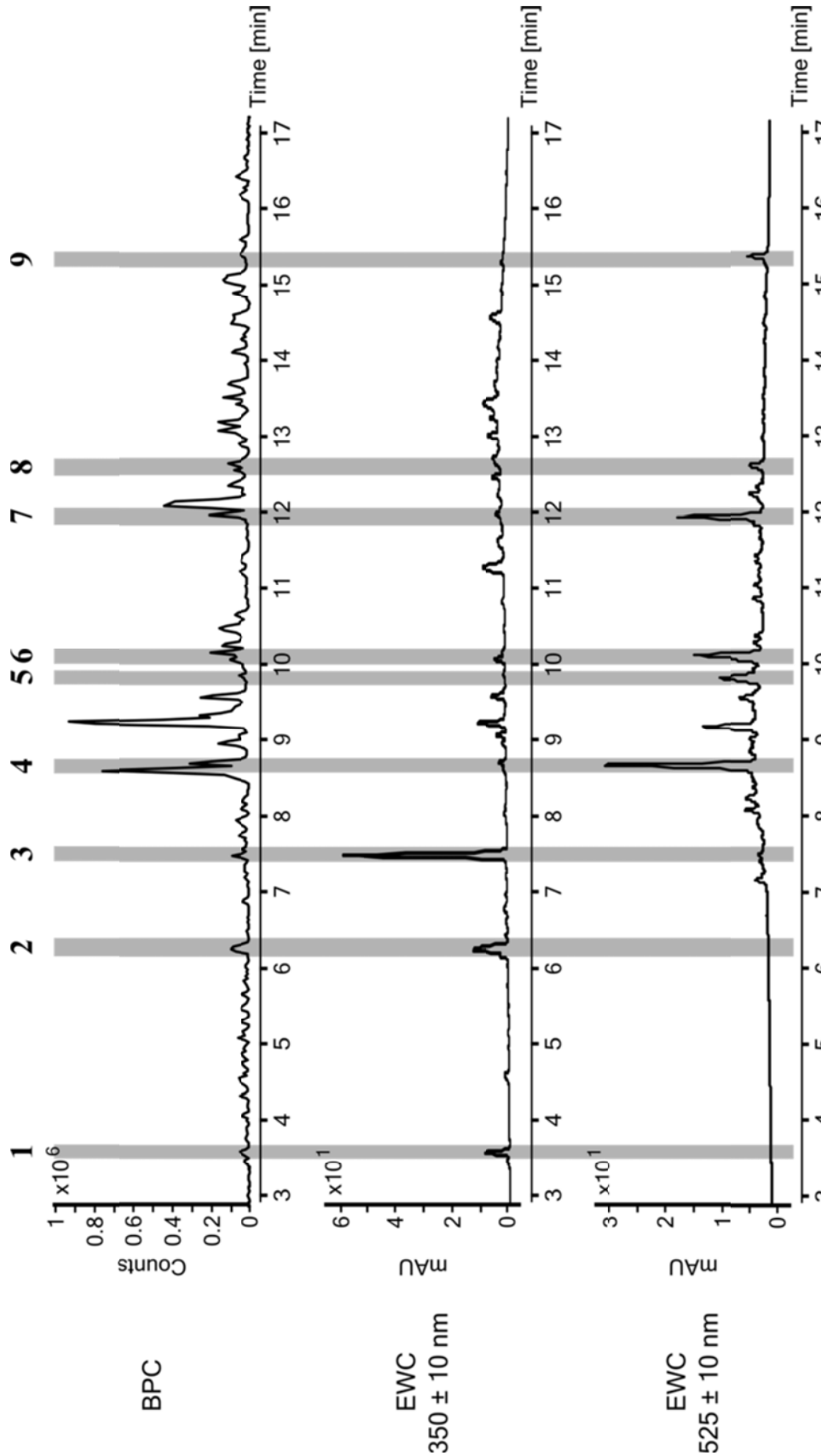


Figure 4.3. Chromatograms of micro-extraction from *T. atroroseus* **IBT11181 grown on CYA for 7 days at 30 °C.** The highlighted compounds are the azaphilone compounds which were linked to their masses and the dereplication and identification of these compounds is further specified in Figure 5.6. However several minor compounds showing the characteristic UV/VIS spectra of monascorubramine were seen but I were unable to link them to their masses. Top: BPC chromatogram, Middle: Extracted wavelength chromatogram, 390 nm ± 10 nm, Bottom: Extracted wavelength chromatogram, 525 nm ± 10 nm.

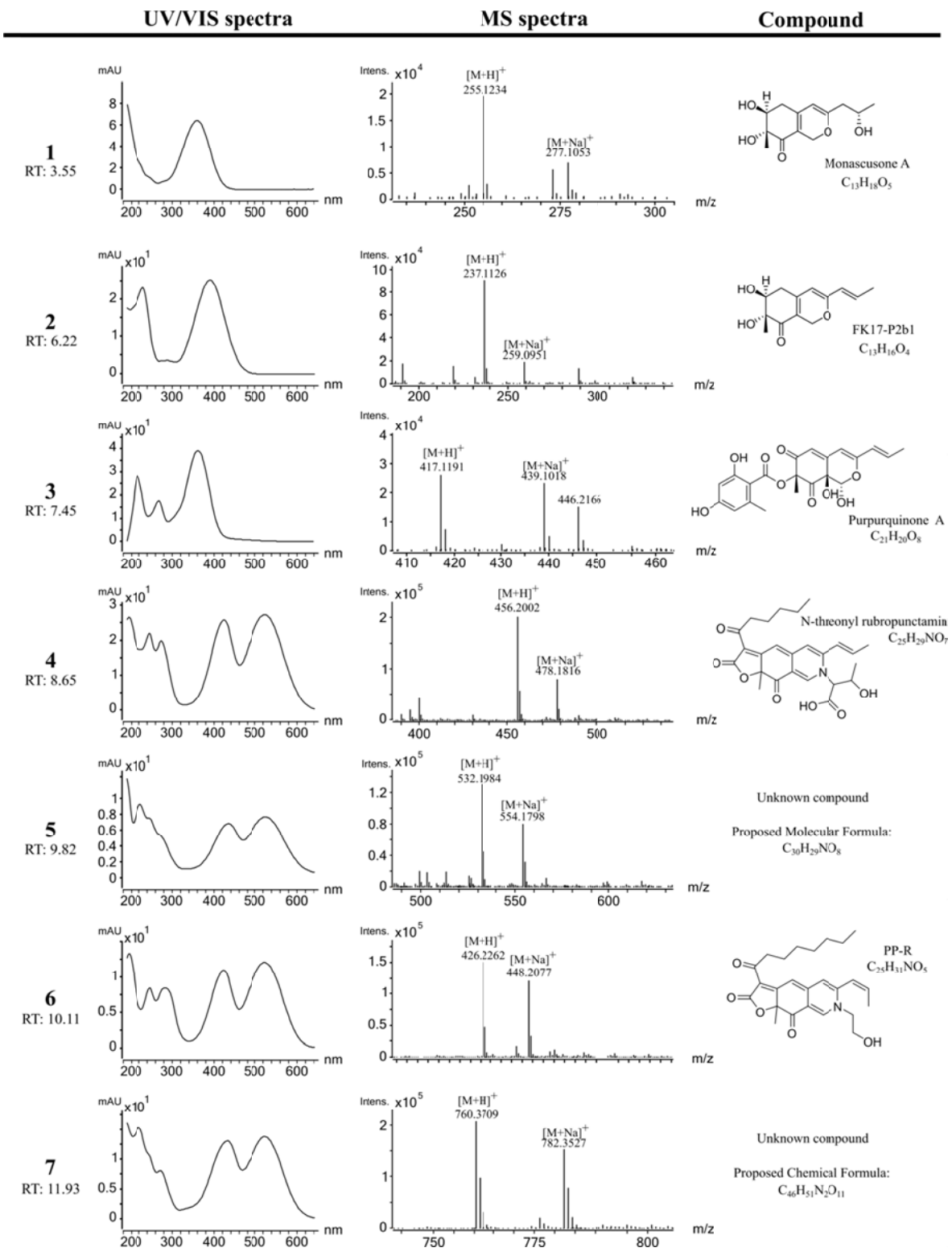


Figure 4.4 (*continues*)

… continued

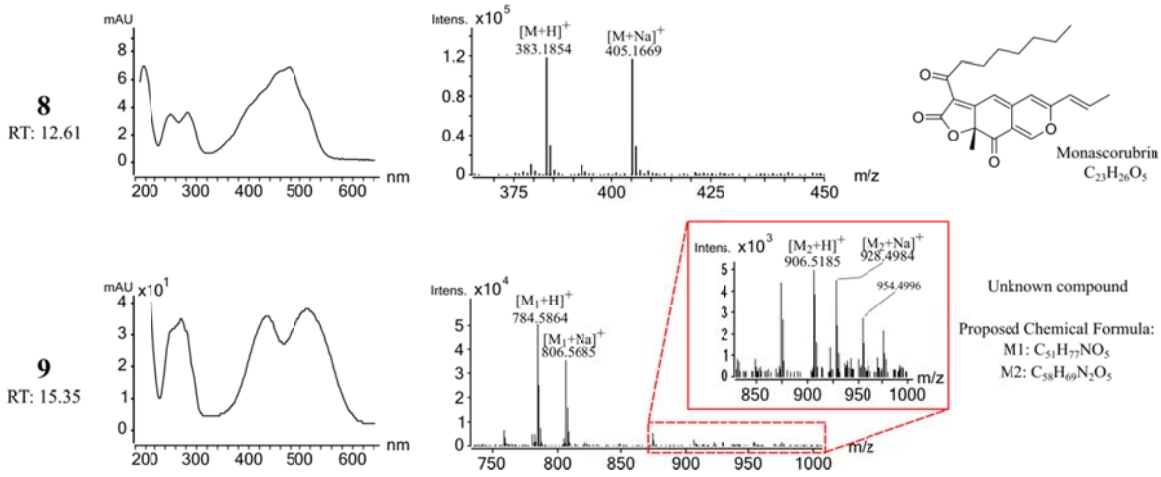


Figure 4.4. Dereplication of known and putatively unknown azaphilones produced by *Talaromyces atroroseus* on CYA. Identification as azaphilones is based on the unique fingerprint UV/VIS-spectra from known azaphilones. Bold numbers refer to the compound numbers in Figure 4.3. For the compounds which were not directly linked to a known azaphilone the proposed molecular formula is stated. See main text for further discussion.

By looking towards alternative masses consistent with the EWC_{525nm} peak I identified an alternative mass via $[M_2+H]^+$=906.5185 supported by the sodium adduct $[M_2+Na]^+$=928.4984. The corresponding calculated formula was $C_{58}H_{49}N_2O_5$. Both dereplicated masses will hence be treated as putatively belonging to a red *Monascus* pigment.

It should be noted that the linking to known compounds are based on the dereplicated masses and UV/VIS spectra since no authentic standards for most of the compounds were available. Only the production of monascorubrin was confirmed by an authentic standard from DTU Systems Biology in-house standard collection.

In summary I confirmed that *T. atroroseus* IBT11181 is capable of producing azaphilones from both the azaphilone families previously identified and linked to their gene clusters in *Talaromyces marneffei* (Woo et al., 2012; Woo et al., 2014). Furthermore I note that MP compounds with both known 3-oxoacyl chain lengths are produced by *T. atroroseus* IBT11181.

4.3.3 Growth and pigment production kinetics

Growth kinetics and total red pigment production of the reactor cultivations are presented in Figure 4.5. The growth curves show a long lag phase of around 30-35 hours with reactor 1 showing a bit longer lag phase resulting in delay of both the exponential phase and the pigment production. The lag phase is followed by an exponential phase before the biomass

reaches its maximum after ~60-70 hours of cultivation. Maximum dry weight concentration was determined to 6.36 ± 0.25 $g.L^{-1}$.

The specific growth rate was calculated to 0.109 ± 0.003 h^{-1} (mean ± standard deviation of duplicates). To my knowledge the only specific growth rate reported in literature of a species potential being *T. atroroseus* or closely related is the growth rate of *Penicillum* sp. HKUCC 8070. Specific growth rates between 0.043 h^{-1} and 0.045 h^{-1} was observed in shake flasks cultivations using complex potato dextrose media, malt extract media as well as in a defined medium based on glucose and glutamate as carbon- and nitrogen-source (Jiang et al., 2005). The specific growth rate for *T. atroroseus* IBT11181 found in the present study is significantly higher. However *T. atroroseus* IBT11181 at least with the given culture and media conditions grows significantly slower than well-established industrial production organisms as *Aspergillus niger* (μ_{max}= 0.22-0.23 h^{-1} (Andersen et al., 2009; Poulsen et al., 2012)) and *Penicillium chrysogenum* (μ_{max} = 0.14-0.15 h^{-1} (Thykaer et al., 2008; Bajaj et al., 2014)).

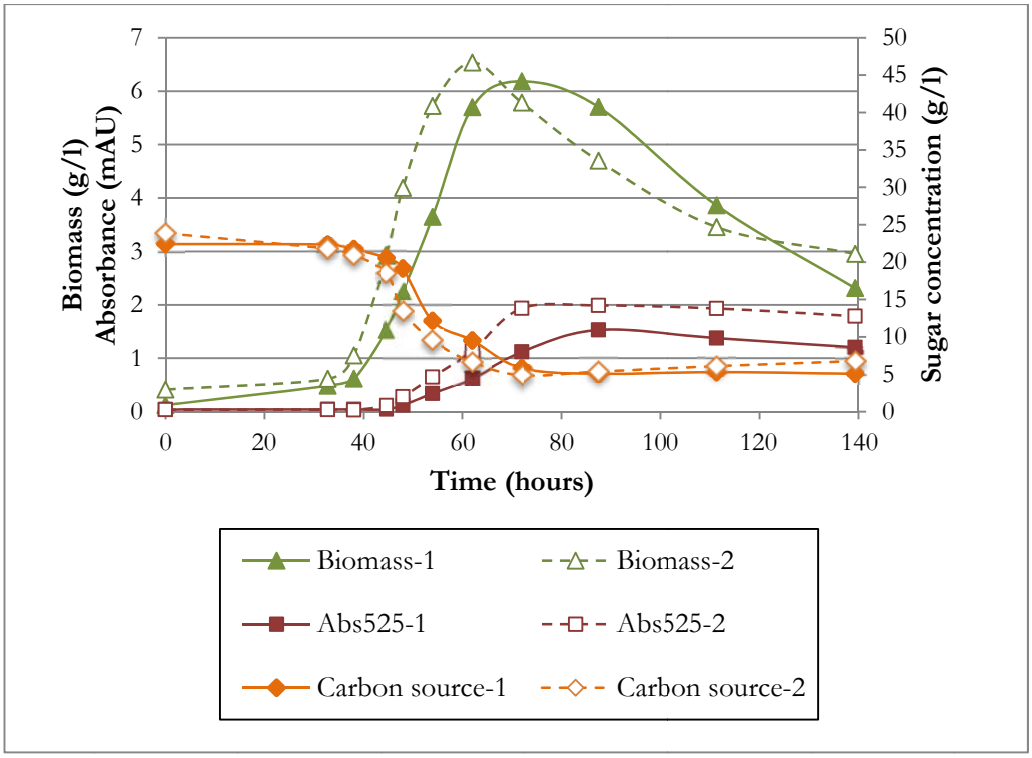


Figure 4.5. Growth kinetics and red pigment production of *Talaromyces atroroseus* IBT11181 on defined media. The concentration of carbon source is the sum of sucrose, glucose, and fructose. Concentration of biomass as well as the production of red pigment measured by the absorbance at 525 nm is given on the left y-axis. Concentration of carbon source is given on the right y-axis.

Red pigment production was measured as the absorbance at 525 nm. The pigment production seems to be growth associated, however it starts a bit later than the biomass production at around 40-45 hours into the fermentation. Maximum pigment produced also reaches it maximum later than the maximum biomass. Based on this the optimal time for termination of fermentation is after around 80 hours, since the total pigment starts to drop hereafter. The decrease of total red pigments is believed to be caused by a combination of breakdown of the pigment by the fungus itself and breakdown due to light exposure, since the reactors were not shielded against light.

4.3.4 Assessment of *Monascus* pigment production by potential *Talaromyces atroroseus* strains in literature

An important aspect of this study was to enhance the resolution of the MP production by a red pigment producing *Talaromyces*. Studies on MP production by *Talaromyces* species focus on evaluating the production potential only measuring the overall absorbance at around 525 nm (Santos-Ebinuma et al., 2013b). However these studies fail to dig a bit deeper and try to identify individual MPs produced.

Furthermore several studies use absorbance values at around 400 nm and 470 nm for measurement of amount of yellow and orange MPs produced (Santos-Ebinuma et al., 2013a; Santos-Ebinuma et al., 2014). However using absorbance at 400 nm and 470 nm for measuring yellow and orange MP production is missing an important aspect. The red MPs have absorption maxima around 525 nm but also have significant absorption from around 390 nm with its local absorption maxima around 425 nm. Thus with co-production of red MPs the wave lengths 400 nm and 470 nm are not suitable for measuring yellow and orange compounds in a complex MP mixture because of the background absorption from red MPs and potentially also from other compounds. To use absorption at wave length around 400 nm and 470 nm for yellow and orange MP production these need to be separated from the red pigments for reliable conclusions.

To try to address both above points fermentation samples were separated by UHPLC (Figure 4.6) allowing for identification (Section 4.3.5) as well as quantitation (Section 4.3.6) of individual compounds using specific absorbance values for yellow, orange and red MPs.

4.3.5 Pigment profile in submerged fermentation

From UHPLC-HRMS analysis of samples from the batch fermentation it was evident that the range of pigments produced was higher than what was observed from the solid media (See Figure 4.7). Masses and chemical formulas were assigned to compounds using the same process as presented above for the solid media. For the red MPs only the 12 largest compounds were considered. The results are summarised in Table 4.1, where also some red MP compounds of which it was not possible to assign masses and formula appear. Both compound **1** and **2** the two putative intermediates in the MP biosynthesis are also found to be produced in the batch fermentations. For most of the red MPs it was not possible to link the proposed molecular formula to any already known compounds, in fact the only known red MP produced and identified was N-threonyl monascorubramine which was also seen produced on solid CYA media. Compound **5** is the only other red MP that for sure is produced on both CYA and in the liquid cultivation.

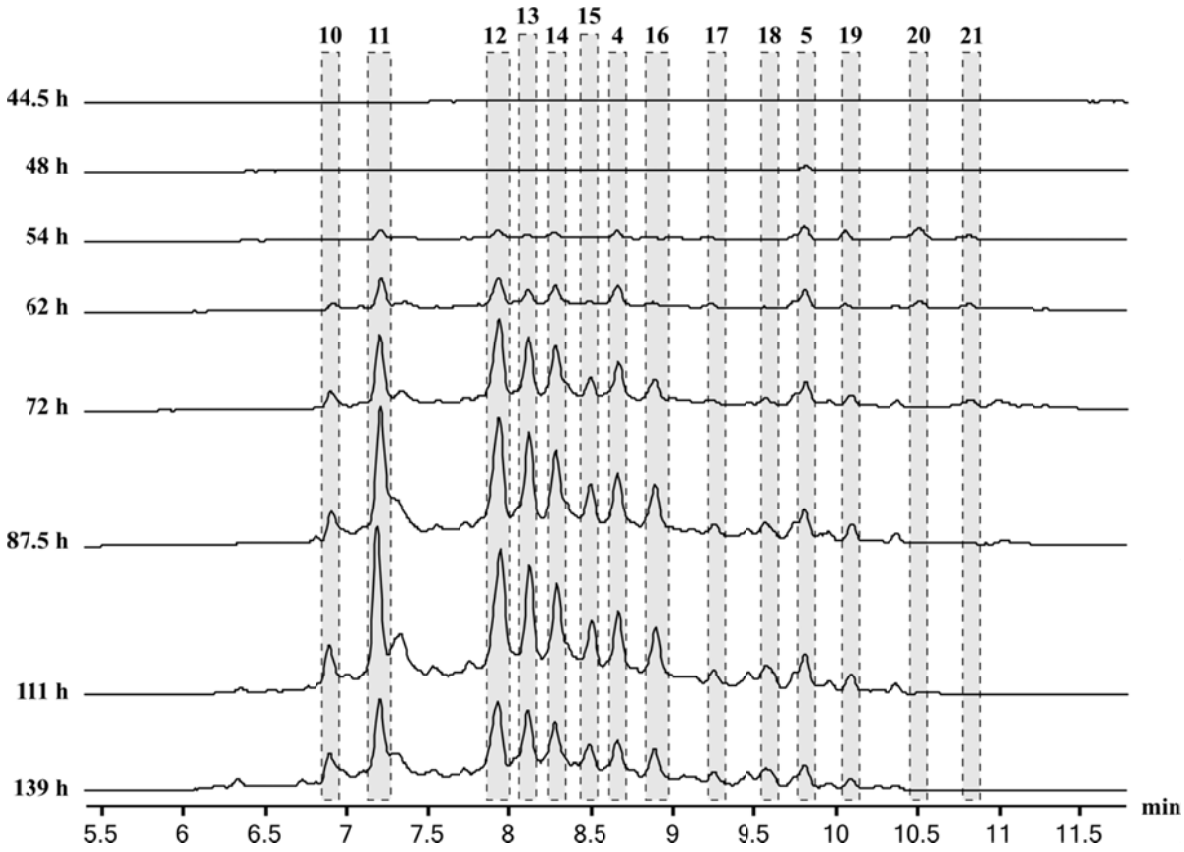


Figure 4.6. Red pigment production profiles from batch fermentation at different time points. Profiles are extracted wavelength chromatograms at 525 nm showing the production of individual pigments during batch fermentation. The shown profiles are from reactor 1. The equivalent profiles for reactor 2 can be found in Appendix D.

Table 4.1. Major azaphilone compounds produced by *T. atroroseus* IBT11181 by batch fermentation.

#	RT (min)	Observed[a] $[M+H]^+$	Proposed molecular formula	Error (ppm)	MP chromo-phore[b]	Proposed compound
1	3.5	255.1230	$C_{13}H_{18}O_5$	-1.18	Yellow[c]	Monascusone A
2	6.2	237.1124	$C_{13}H_{16}O_4$	-1.12	Yellow	FK17-P2b1
10	6.9	n.a.			Red	-
11	7.2	610.2048	$C_{35}H_{31}N_2O_9$	3.87	Red	Unknown
12	7.9	541.2175	$C_{28}H_{32}N_2O_9$	1.03	Red	Unknown
13	8.1	528.1862	$C_{27}H_{29}NO_{10}$	0.42	Red	Unknown
14	8.3	542.2015	$C_{28}H_{31}NO_{10}$	1.06	Red	Unknown
15	8.5	514.2065	$C_{27}H_{31}NO_9$	1.28	Red	Unknown
4	8.7	456.2020	$C_{25}H_{29}NO_7$	-0.71	Red	N-threonyl rubropunctamin
16	8.9	576.2219	$C_{32}H_{33}NO_9$	1.58	Red	Unknown
17	9.2	n.a.			Red	-
18	9.6	n.a.			Red	-
5	9.8	532.1960	$C_{30}H_{29}NO_8$	1.1	Red	Unknown
19	10.1	n.a.			Red	-
20	10.5	413.1591	$C_{23}H_{24}O_7$	0.92	Orange	PP-O
21	10.8	413.1591	$C_{23}H_{24}O_7$	0.92	Orange	PP-O

n.a. = not assigned; it was not possible to assign masses to these peaks from the EWC at 525 nm.
[a]All observed $[M+H]^+$ ions were supported by $[M+Na]^+$ with an error below 3 ppm
[b]UV/VIS spectra for the compounds showing the typical orange or red *Monascus* pigment chromophore can be found in Appendix C.
[c]This yellow UV/VIS spectra is different from the UV/VIS spectra related to the yellow MP chromophore seen in monascorubrin and rubropunctatin

Of the unknown red compounds showing the typical red MP UV/VIS spectrum the proposed molecular formulas is different from the molecular formulas from the original red MPs monascorubramine and rubropunctamin. The molecular differences between the two original MPs and the unknown red MPs produced by *T. atroroseus* during the fermentation is believed to be due to nitrogen attached amino acids or sugars. However combining the 20 proteogenic amino acids with either monascorubramine or rubropunctamine fits the proposed molecular formulas. Comparing the proposed molecular formulas for the unknown MPs with the two original MPs especially the difference in number of oxygen atoms is apparent. While monascorubramine and rubropunctamine both contain four oxygen the proposed formulas from the unknown MPs contain between eight and ten oxygen suggesting that these are largely oxidized compared to the original red MPs. Furthermore it is noted that compound **11** and **12** both contain two nitrogen as opposed to one observed in other MPs. This might be the result of an N-linked amino acid with site chain containing one nitrogen, i.e. asparagine, glutamine, lysine or tryptophan. However the molecular formula of monascorubramine or rubropunctamine linked with any of these amino acids does not match the proposed molecular formula of **11** or **12**. As an example is the molecular formula of N-glutamyl monascorubramine $C_{28}H_{34}N_2O_7$ which differs from **12** by having two oxygens exchanged for

two hydrogens. This equals a methyl group fully oxidized to a carboxylic acid. This will be further discussed in Section 4.3.7.

Apart from the red MPs two MPs with identical masses and exhibiting the orange MP chromophore were produced. These matches the orange PP-O which has previously been identified produced by *T. atroroseus* (Ogihara & Oishi, 2002). It should be noted that only the filtrated fermentation broth and not the biomass from the liquid cultivation that was analysed for pigments production. Thus hence only extracellular pigment was detected and more hydrophobic pigments such as **8** and **9** might still have been produced but kept in the mycelia and has hence not been detected in the submerged fermentation.

4.3.6 Time-course comparison of pigment profiles

One of the purposes of this study was to address the dynamic production of secreted MPs by *T. atroroseus*. Especially addressing whether different MPs are produced at different time of the cultivation is a focus area. This was addressed by comparing semi-quantitative production profiles of individual MPs.

With no authentic standard available individual pigments were quantified as their peak heights in extracted wave length chromatograms corresponding to yellow, orange and red MPs as specified in *Materials and methods*. The time-course pigment production of the individual MPS is shown in Figure 4.6. For the two yellow putative intermediates in the MP biosynthesis compound **1** and **2** the trend is that they follow the exponential growth with compound **1** stabilizing after ~50 hours, while compound **2** first stabilizes after ~80 hours.

For the orange MPs compound **20** and **21** they are both rapidly produced and secreted from around 40 hours into the cultivation. However after reaching their maximum concentration they are rapidly being decreased. This suggests that they are converted into red MPs by reaction with free NH_3 or amino acids. This is in accordance with the reaction generally associated with the azaphilone class of compounds creating vinylogous γ-pyridones (Stadler et al., 1995; Gao et al., 2013). It is furthermore worth to note that the amount of orange MPs produced is minor compared to the amount of red MPs (Figure 4.6-4.7).

Looking towards the red MPs four of the pigments are the main contributors to the overall red production, i.e. compound **11-14** (Figure 4.6-4.7). The production profiles of all the red pigments are however very similar suggesting that these are not interconverted to each other and are rather produced simultaneously with some of the red MPs being produced at a higher rate than others.

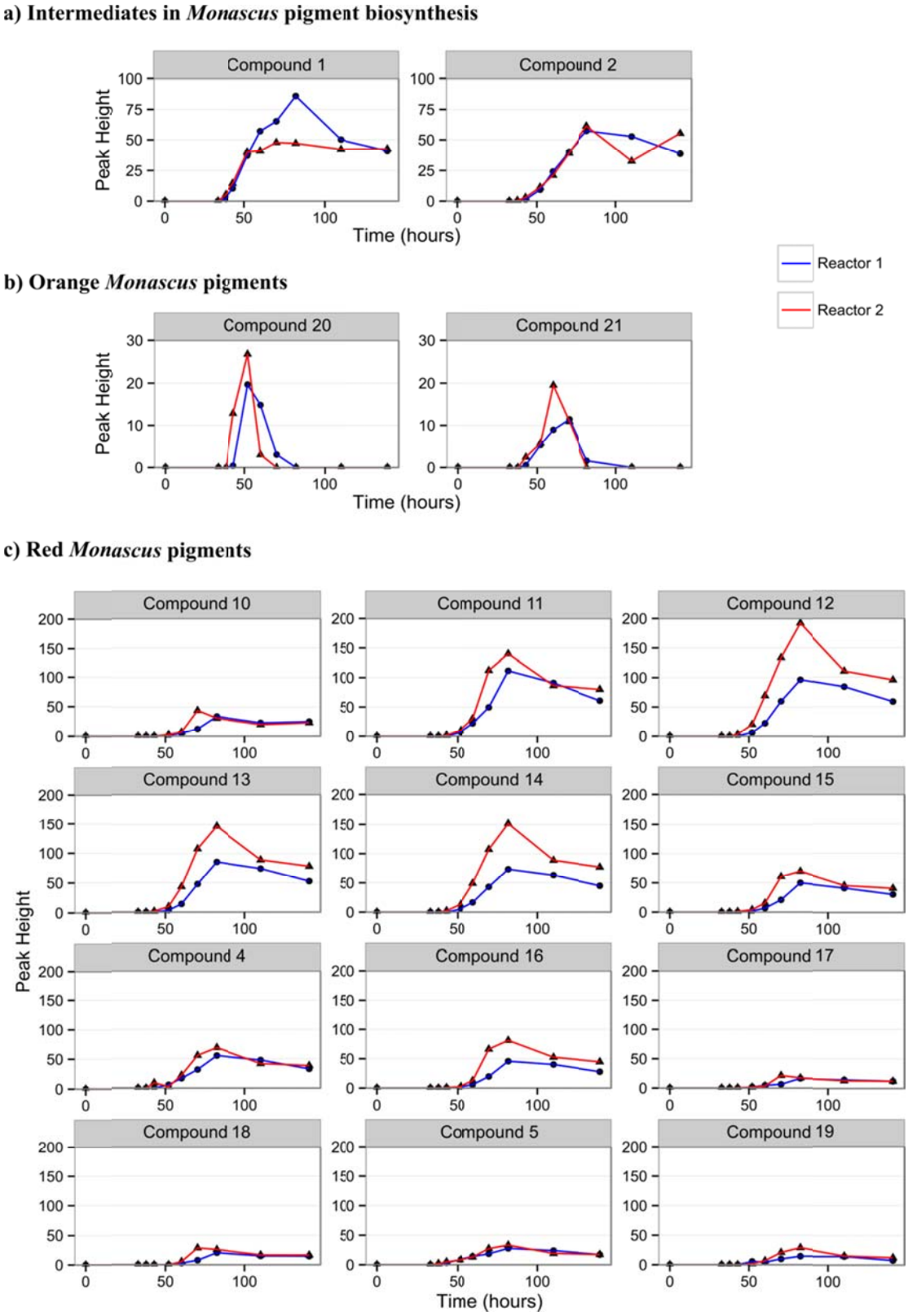


Figure 4.7. Production profiles of individual *Monascus* pigments from batch fermentations. Note that the scale of the Y-axis is different between the three panels, so the produced amount cannot be directly compared between panels, but only within panels. Panel a) Production of two intermediates in the MP biosynthesis. Panel b) Production of two orange MPs. Panel c) Production profiles of 12 red nitrogen-containing MPs. The production profiles are ordered according to the retention times rather than the compound numbering.

4.3.7 Novel *Monascus* pigment

The major pigment, compound **12** was purified and structure elucidated by NMR (See appendix D purification and structure elucidation). The compound turned out to be the glutamine derivative of PP-V (Ogihara et al., 2000), see Figure 4.8. This is a novel variation of the expanding family of MPs. PP-V and **12** both contains carboxyl group at the end of the polyketide chain (Figure 4.8). This has to my knowledge not been observed in any other MPs, but a similar carboxylation is found in mitorubrinic acid (Natsume et al., 1985) also produced by *T. atroroseus* species (Chapter 3 and Frisvad et al., 2013). The finding of the carboxyl group confirms my previous structural speculations for compound **12**. Compound **12** will be called N-glutamyl monascorubraminic acid reflecting the nitrogen linked glutamine as well as the carboxylic acid form of monascorubramin.

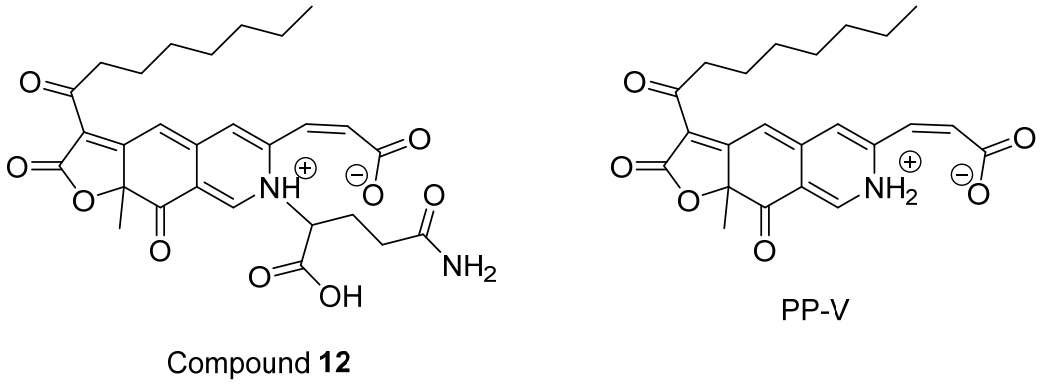


Figure 4.8. Structure of compound 12 as elucidated by NMR. For specific detail on the structure elucidation see appendix D.

The observed carboxylation in the MP compounds produced by *T. atroroseus* could easily also be present in several of the unknown red MPs observed in the fermentation broth. With the assumption that the unknown MPs have a carboxyl group at the polyketide end and changing the 3-oxoacyl chain length as well as the N-linked amino acid probable structures for several of the unknown MPs from Table 4.1 is hypothesized (Figure 4.9). Specifically compound **13**, **14**, **15**, **16** match with the potential structures of the acid derivatives of monascorubramine linked with aspartic acid, glutamic acid, threonine and tyrosine respectively. Furthermore applying the same assumption of carboxylation of the polyketide end to compound **4**, which was dereplicated as N-threonyl rubronpunctamine above, it is possible that this compound is not N-threonyl monascorubramine, but rather an acid form of PP-R.

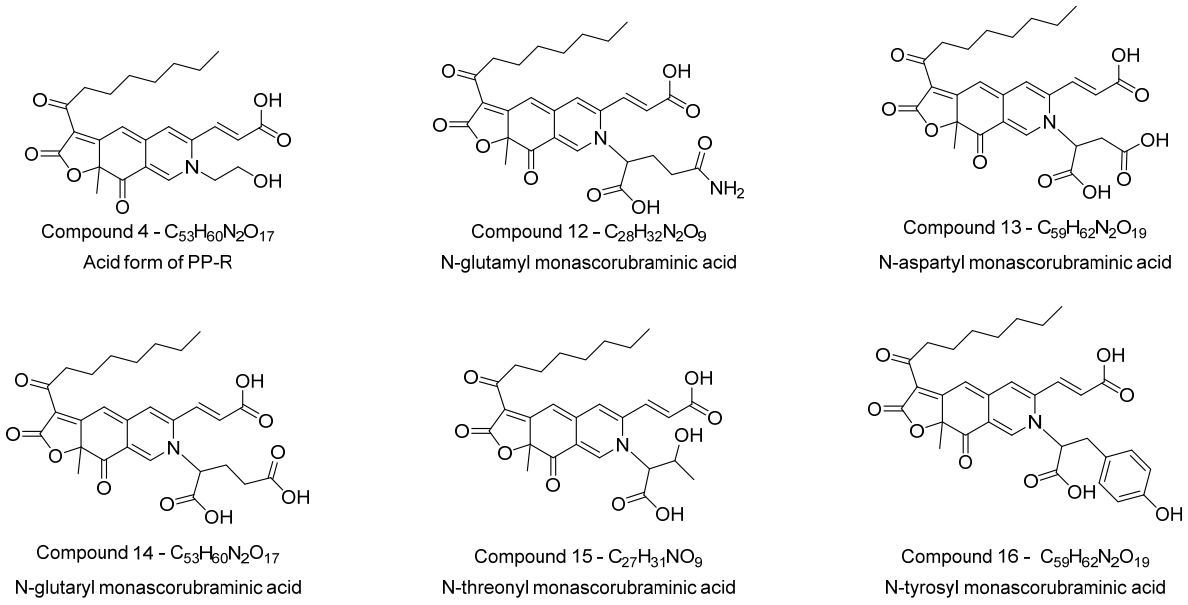


Figure 4.9. Hypothetical structures of some of the unknown red MPs as well as the novel structure elucidated MP N-glutamyl monascorubraminic acid. Adding the criterion of a carboxylic acid residue at the polyketide chain end as seen in the novel N-glutamyl monascorubraminic acid resulted in hypothetical structures for some of the unknown red MPs as well as an alternative structure of compound 4 which was previously dereplicated to N-threonyl rubropunctamin.

4.3.8 Future perspectives for quantification of *Monascus* pigments

The presented quantification of MPs can be subject to improvement since the unit of peak height are not suitable for comparison between studies. One future aspect for improving this will be by using authentic standards to convert the peak heights to a quantitative measurement in $g.L^{-1}$. The optimal would be to have authentic standards for every produced MP to use for creating external calibration curves. However with the broad range of MPs produced it would be labour intensive to purify enough of every MP to create reliable calibration curves. Instead I would suggest purifying 2-3 compounds in appropriate amounts for creating calibration curves. These calibration curves could then be used to convert the peak heights to number of red MP chromophores. I believe that due to the conserved MP chromophore across all red MP pigments the peak height of the individual compounds in an EWC_{525nm} could be used for reliable quantification of the number of chromophores using calibration curves of one or more of the other red MPs. However this of course needs to be validated. With the MS identified compound formulas it is the possible to convert the number of chromophores to masses.

4.4 Conclusion

In the present study it is shown that fermentative production of MPs by *T. atroroseus* IBT11181 in bioreactors under controlled conditions is reproducible. Under the given conditions a broad range of red MPs were produced, many of these being putatively novel compounds. During the fermentation *T. atroroseus* also produces and secrets small amounts of two orange MPs. These however disappear from the fermentation supernatant and have probably reacted with free amino acids to create varieties of red MPs. One of the produced *Monascus* pigment has been isolated and structure elucidated to be a monascorubramine analog with a carboxylic acid at the polyketide end as also seen in PP-V and furthermore linked to glutamine. Based on the structure hypothetical structures for several of the other unknown MPs is suggested. Further purification studies will confirm or refute these hypothetical structures.

Acknowledgements

Alexander Rosenkjær is thanked for technical assistance in running the sugar quantifications.

4.5 References

Andersen, M.R., Lehmann, L. & Nielsen, J., **2009**. Systemic analysis of the response of *Aspergillus niger* to ambient pH. *Genome Biology*, **10**: R47. doi:10.1186/gb-2009-10-5-r47

Arai, T. & Sano, H., **1994**. Novel UV-absorbing compounds FK17-P2a, FK17-P2b1, FK17-P2b2, and FK17-P3 and manufacture of the compounds with *Aspergillus sp.* (in Japanese). JP 6329576.

Bajaj, I., Veiga, T., van Dissel, D., Pronk, J.T. & Daran, J.-M., **2014**. Functional characterization of a *Penicillium chrysogenum* mutanase gene induced upon co-cultivation with *Bacillus subtilis*. *BMC Microbiology*, **14**: 114. doi:10.1186/1471-2180-14-114

Bijinu, B., Suh, J.-W., Park, S.-H. & Kwon, H.-J., **2014**. Delineating *Monascus* azaphilone pigment biosynthesis: oxidoreductive modifications determine the ring cyclization pattern in azaphilone biosynthesis. *RSC Advances*, **4**: 59405–59408. doi:10.1039/C4RA11713A

Espinoza-Hernandez, T.C., Rodriguez-Herrera, R., Aguilar-Gonzalez, C.N., Lara-Victoriano, F., Reyes-Valdes, M.H., et al., **2013**. Characterization of three novel pigment-producing *Penicillium* strains isolated from the Mexican semi-desert. *African Journal of Biotechnology*, **12**: 3405–3413. doi:10.5897/AJB2013.12338

Frisvad, J.C., Filtenborg, O., Samson, R.A. & Stolk, A.C., **1990**. Chemotaxonomy of the genus *Talaromyces*. *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology*, **57**: 179–189. doi:10.1007/BF00403953

Frisvad, J.C., Yilmaz, N., Thrane, U., Rasmussen, K.B., Houbraken, J., et al., **2013**. *Talaromyces atroroseus*, a new species efficiently producing industrially relevant red pigments. *PloS ONE*, **8**: e84102. doi:10.1371/journal.pone.0084102

Gao, J.-M., Yang, S.-X. & Qin, J.-C., **2013**. Azaphilones: chemistry and biology. *Chemical Reviews*, **113**: 4755–4811. doi:10.1021/cr300402y

Jiang, Y., Li, H., Chen, F. & Hyde, K., **2005**. Production potential of water-soluble *Monascus* red pigment by a newly isolated *Penicillium sp. Journal of Agricultural Technology*, 113–126.

Jongrungruangchok, S., Kittakoop, P., Yongsmith, B., Bavovada, R., Tanasupawat, S., et al., **2004**. Azaphilone pigments from a yellow mutant of the fungus *Monascus kaoliang*. *Phytochemistry*, **65**: 2569–75. doi:10.1016/j.phytochem.2004.08.032

Mapari, S.A.S., **2008**. *Chemotaxonomic exploration of fungal biodiversity for polyketide natural food colorants*. PhD thesis, Technical University of Denmark.

Mapari, S.A.S., Meyer, A.S., Frisvad, J.C. & Thrane, U., **2009a**. Production of *Monascus*-like azaphilone pigments. Patent No. WO2009/026923 A2

Mapari, S.A.S., Meyer, A.S., Thrane, U. & Frisvad, J.C., **2009b**. Identification of potentially safe promising fungal cell factories for the production of polyketide natural food colorants using chemotaxonomic rationale. *Microbial cell factories*, **8**: 24. doi:10.1186/1475-2859-8-24

Méndez, A., Pérez, C., Montañéz, J.C., Martínez, G., Aguilar, C.N.C.N., et al., **2011**. Red pigment production by *Penicillium purpurogenum* GH2 is influenced by pH and temperature. *Journal of Zhejiang University. Science. B*, **12**: 961–8. doi:10.1631/jzus.B1100039

Mendez-Zavala, A., Contreras-Esquivel, J.C., Lara-Victoriano, F., Rodriguez-Herrera, R. & Aguilar, C.N., **2007**. Fungal production of the red pigment using a xerophilic strain *Penicillium purpurogenum* GH-2. *Revista Mexicana De Ingenieria Quimica*, **6**: 267–273.

Ming Ge, H., Yun Zhang, W., Ding, G., Saparpakorn, P., Chun Song, Y., et al., **2008**. Chaetoglobins A and B, two unusual alkaloids from endophytic *Chaetomium globosum* culture. *Chemical Communications (Cambridge, England)*, 5978–80. doi:10.1039/b812144c

Natsume, M., Takahashi, Y. & Marumo, S., **1985**. (-)-Mitorubrinic acid, a morphogenic substance inducing chlamydospore-like cells, and its related new metabolite, (+)-mitorubrinic acid-B, isolated from *Penicillium funiculosum*. *Agricultural and Biological Chemistry*, **49**: 2517–2519.

Ogihara, J., Kato, J., Oishi, K. & Fujimoto, Y., **2001**. PP-R, 7-(2-hydroxyethyl)-monascorubramine, a red pigment produced in the mycelia of *Penicillium* sp. AZ. *Journal of Bioscience and Bioengineering*, **91**: 44–47.

Ogihara, J., Kato, J., Oishi, K., Fujimoto, Y. & Eguchi, T., **2000**. Production and structural analysis of PP-V, a homologue of monascorubramine, produced by a new isolate of *Penicillium* sp. *Journal of Bioscience and Bioengineering*, **90**: 549–554.

Ogihara, J. & Oishi, K., **2002**. Effect of ammonium nitrate on the production of PP-V and monascorubrin homologues by *Penicillium* sp. AZ. *Journal of Bioscience and Bioengineering*, **93**: 54–59.

Poulsen, L., Andersen, M.R., Lantz, A.E. & Thykaer, J., **2012**. Identification of a transcription factor controlling pH-dependent organic acid response in *Aspergillus niger*. *PLoS ONE*, **7**: e50596. doi:10.1371/journal.pone.0050596

Samson, R.A., Houbraken, J., Thrane, U., Frisvad, J.C. & Andersen, B., **2010**. *Food and indoor fungi*, CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands.

Samson, R.A., Yilmaz, N., Houbraken, J., Spierenburg, H., Seifert, K.A., et al., **2011**. Phylogeny and nomenclature of the genus *Talaromyces* and taxa accommodated in *Penicillium* subgenus *Biverticillium*. *Studies in Mycology*, **70**: 159–183. doi:10.3114/sim.2011.70.04

Santos-Ebinuma, V.C., Roberto, I.C., Francisca, M., Teixeira, M.F.S, Jr, A.P., et al., **2014**. Improvement of submerged culture conditions to produce colorants by Penicillium purpurogenum. , **742**: 731–742.

Santos-Ebinuma, V.C., Roberto, I.C., Teixeira, M.F.S. & Jr., A.P., **2013a**. Improving of red colorants production by a new *Penicillium purpurogenum* strain in submerged culture and the effect of different parameters in their stability. *Biotechnology Progress*, **29**: 778–785.

Santos-Ebinuma, V.C., Teixeira, M.F.S. & Jr., A.P., **2013b**. Submerged culture conditions for the production of alternative natural colorants by a new isolated *Penicillium purpurogenum* DPUA 1275. *Journal of Microbiology and Biotechnology*, **23**: 802–810.

Stadler, M., Anke, H., Dekermendjian, K., Reiss, R., Sterner, O., et al., **1995**. Novel bioactive azaphilones from fruit bodies and mycelial cultures of the ascomycete *Bulgaria inquinans* (Fr.). *Natural Product Letters*, **7**: 7–14.

Tabata, Y., Ikegami, S., Yaguchi, T., Sasaki, T., Hoshiko, S., et al., **1999**. Diazaphilonic acid, a new azaphilone with telomerase inhibitory activity. *The Journal of Antibiotics*, **52**: 412–414.

Thykaer, J., Rueksomtawin, K., Noorman, H. & Nielsen, J., **2008**. NADPH-depandent glutamate dehydrogenase in *Penicillium chrysogenum* is involved in regulation of b-lactam production. *Microbiology*, **154**: 1242–1250. doi:10.1099/mic.0.2007/010017-0

Wang, H., Wang, Y., Wang, W., Fu, P., Liu, P., et al., **2011**. Anti-influenza virus polyketides from the acid-tolerant fungus *Penicillium purpurogenum* JS03-21. *Journal of Natural Products*, **74**: 2014–2018. doi:10.1021/np2004769

Woo, P.C.Y., Lam, C.-W., Tam, E.W.T., Lee, K.-C., Yung, K.K.Y., et al., **2014**. The biosynthetic pathway for a thousand-year-old natural food colorant and citrinin in *Penicillium marneffei*. *Scientific Reports*, **4**: 6728. doi:10.1038/srep06728

Woo, P.C.Y., Lam, C.-W., Tam, E.W.T., Leung, C.K.F., Wong, S.S.Y., et al., **2012**. First discovery of two polyketide synthase genes for mitorubrinic acid and mitorubrinol yellow pigment biosynthesis and implications in virulence of *Penicillium marneffei*. *PLoS Neglected Tropical Diseases*, 6: e1871. doi:10.1371/journal.pntd.0001871

Yilmaz, N., **2015**. *Employing a polyphasic taxonomy in Talaromyces*. PhD thesis. Utrecht University.

Chapter 5

Missing *Monascus* pigment polyketide synthase in *Talaromyces atroroseus* IBT11181

An important aim of the presented PhD study was to investigate the genetic basis of *Monascus* pigment production in *Talaromyces* species. To facilitate this *Talaromyces atroroseus* IBT11181 was genome sequenced to serve as model organism for studies of both biosynthesis and fermentative production of *Monascus* pigments.

In this chapter I present the genome of *Talaromyces atroroseus* IBT11181, the first *Talaromyces* genome of one of the excellent *Monascus* pigment producers. Orthology comparison of PKS genes within *Talaromyces* revealed that the *Monascus* pigment gene cluster lacks the known *Monascus* pigment PKS. I show that the NR-PKS responsible for producing the mitorubrins also is responsible for the production of the *Monascus* pigments in *Talaromyces atroroseus*. This is to my knowledge the first example of a fungal polyketide gene cluster, which is missing the PKS gene but remains functional due to overlapping biosynthetic pathway with another polyketide gene cluster.

This chapter is together with Chapter 6 in the process of being condensed into a research manuscript.

5.1 Introduction

The pigment producing ability of selected *Talaromyces* species and especially the efficient secretion of red *Monascus* pigments (MPs) make these species interesting mycotoxin-free alternatives to *Monascus* species for industrial scale production of MPs (Mapari et al., 2008; Frisvad et al., 2013; Dufossé et al., 2014). Currently several investigations of the cellular performance and pigment producing potential of various *T. atroroseus* strains are being carried out (Mendez-Zavala et al., 2007; Méndez et al., 2011; Espinoza-Hernández et al., 2013; Santos-Ebinuma et al., 2013a; Santos-Ebinuma et al., 2013b), including ongoing studies at Section for Eukaryotic Biotechnology, DTU Systems Biology (Chapter 4; Nymschefsky, unpublished).

At the initiation of this study the genetic basis for MP production was not known. At the same time no *Monascus* genome was public available for possible identification of the MP gene cluster. Among *Talaromyces* genomes *Talaromyces stipitatus* and *T. marneffei* were public available, however *T. stipitatus* has never been reported to produce MPs, and since *T. marneffei* is an opportunistic human pathogen it is not the obvious choice to study the MP cluster in the *Talaromyces* genus. In a 2013 study by Arai *et al.* (Arai et al., 2013) 454 sequencing of *T. atroroseus* IAM15392 (Published as *P. purpurogenum*) was used as basis for qPCR expression comparison of genes (*glnA*: glutamine synthetase; *gdhA:* glutamate dehydrogenase) involved in the ammonia assimilation pathway in response to media concentration of L-glutamate and L-glutamine. However, the draft sequence data has not been made public available. Thus prior to this study no genome from a suitable MP producing strain was available to serve as model organism for further investigations of the pigment production potential of MP producing *Talaromyces*.

The aim of the present study was to genome sequence a suitable MP producing *Talaromyces* to serve as model genome for both studying the MP biosynthesis as well as provide the basis for further understanding and optimizing *T. atroroseus* as cell factory for MP production. *T. atroroseus* IBT11181 (from now referred to as *T. atroroseus*) has previously shown to be promising as a cell factory for red pigment production and hence was chosen for genome sequencing.

During the process of the genome sequencing of *T. atroroseus* the gene cluster for MP production was reported by Balakrishnan *et al.* (Balakrishnan et al., 2013). By T-DNA random mutagenesis in *Monascus purpureus* they created an albino mutant with a T-DNA insertion

upstream of the transcritpition factor *mppR1*. Following targeted konockout of the clustered PKS gene, *MpPKS5*, verified the gene cluster as responsible for MP production in *M. purpureus*. This observation was repeated by knockout of the PKS gene, *pks3*, from the homologous gene cluster in *T. marneffei* PM1 (Woo et al., 2014).

In the present study the first whole genome of a suitable red pigment producing *Talaromyces* strain is presented. Comparison of its polyketide potential reveals surprisingly that it does not have the MP PKS as found in *T. marneffei* (Woo et al., 2014). Knockout of the PKS from the mitorubrin gene cluster revealed that the MP biosynthesis in *T. atroroseus* is entangled with the biosynthesis of the mitorubrins, which provide new insights into polyketide biosynthesis complexity in filamentous fungi.

5.2 Materials and Methods

5.2.1 Genome sequences

The genomes including annotations of *Talaromyces stipitatus* ATCC10500, *Talaromyces marneffei* PM1 and *Talaromyces marneffei* ATCC18224 were retrieved from GenBank in the form of assembly GCA_000003125.1, GCA_000750115.1 and GCA_000001985.1 respectively. Genomes and annotations of *Talaromyces aculeatus* ATCC10409, *Monascus ruber* NRRL1597 and *Monascus purpureus* NRRL 1596 were retrieved from the JGI Genome Portal; all as version 1.0.

5.2.2 Strain and culture conditions

Talaromyces atroroseus IBT11181 was obtained from the IBT Culture Collection at Department of Systems Biology at Technical University of Denmark. The strain is also deposited in the CBS collection at CBS-KNAW, the Netherlands, as CBS 123796 and CBS 238.95. Fungal media was prepared as described by Samson *et al.* (Samson et al., 2010). CYA agar medium was used for general propagation of *T. atroroseus*.

5.2.3 Extraction of genomic DNA for genome sequencing and southern blot

Genomic DNA was extracted with a slightly modified protocol of the CTAB-method used by Fulton and co-workers (Fulton et al., 1995). Briefly *T. atroroseus* was propagated in three-point colonies on CYA medium for 7 days at 30 °C. Spores from one plate was harvested and

inoculated in a 500 ml shake flask containing 100 ml Czapek yeast extract medium (30 $g.L^{-1}$ sucrose; 3 $g.L^{-1}$ $NaNO_3$; 1 $g.L^{-1}$ K_2PO_4; 0.5 $g.L^{-1}$ KCl; 0.5 $g.L^{-1}$ $MgSO_4$; 0.01 $g.L^{-1}$ $FeSO_4$; 5 $g.L^{-1}$ Yeast extract; 0.01 $g.L^{-1}$ $ZnSO_4 \cdot 7H_2O$; 0.005 $g.L^{-1}$ $CuSO_4 \cdot 5H_2O$) and grown at 30 °C with shaking for 40 hours. Mycelia were harvested in sterile mira-cloth, flash frozen in liquid nitrogen and lyophilized. The lyophilized and frozen mycelia were ground to a fine powder using liquid nitrogen. Lysis buffer was freshly prepared by mixing 3.75 vol Buffer A (0.35 M sorbitol, 0.1 M Tris-HCL, 5 mM EDTA, pH 8), 3.75 vol Buffer B (0.2 M Tris-HCL, 50 mM EDTA, 2 M NaCl, 2 % CTAB, pH 8), 1.5 vol 5% Sarkosyl, 1 vol 1% PVP and Proteinase K (New England Biolabs) was added to a final concentration of 150 μg/ml. The lysis buffer was added to ground fungal mycelium and the mixture was incubated at 65 °C for 30 min. After incubation 1 volume potassium acetate (5 M) was added and the mixture incubated on ice for 30 min followed by 20 min centrifugation at 5,000 x g at 4 °C. The supernatant was then extracted twice with Phenol:Chloroform:Isoamylalcohol (25:24:1) and treated with RNAse A for 2 hours. The gDNA was precipitated with ice-cold ethanol, resuspended in TE buffer and stored at -80 °C until usage.

5.2.4 Genome sequencing and assembly

Preparation of sequencing libraries and whole genome shotgun sequencing was performed by Beijing Genome Institute (BGI), Hong-Kong. Two libraries were constructed; a 180 bp paired end library and a 6 kb mate-pair library, and sequenced on the Illumina HiSeq 2000 Platform. *De novo* assembly of the genome was performed with the ALLPATHS-LG algorithm (Gnerre et al., 2011).

The raw data files have been submitted to the Sequencing Reads Archive (SRA) at National Center for Biotechnology Information (NCBI) under the accessions SRR2071965 and SRR2071966, and will be public available upon publication of the genome.

5.2.5 Gene calling and functional annotation

Repeats in the draft genome were masked using RepeatMasker 4.0.5 (Smit et al., 2014). A *T. atroroseus* specific repeat database was generated *ab initio* using RepeatModeler 1.0.8 (Smit & Hubley, 2014) and used together with Repbase (Jurka et al., 2005) in the repeat masking. Gene calling was performed on the masked genome using three different *ab initio* gene calling algorithms: i) AUGUSTUS 3.0.3 (Stanke et al., 2006; Stanke et al., 2008) trained with

Aspergillus nidulans and *Fusarium graminearum*, ii) FGENESH 3.1.2 (SoftBerry) (Salamov & Solovyev, 2000) trained with *Aspergillus nidulans*, *Fusarium graminearum* and *Penicillium funicolosum* respectively, and iii) Genemark-ES (Ter-Hovhannisyan et al., 2008). The different gene models from the *ab initio* predictions were merged to a consensus gene model using EVidenceModeler (Haas et al., 2008). In order to estimate the completeness of the assembly and gene calling the CEGMA pipeline was used (Parra et al., 2007). It searches for the presence of 248 Core Eukaryotic Genes (CEGs) identified to be present and highly conserved throughout the eukaryotic kingdom (Parra et al., 2009). The pre-computed default cut-off scores were used and genes with alignment length of minimum 70% were counted as complete genes.

The predicted proteins were functionally annotated by BLASTP best hit against UniProtKB/SwissProt (Boeckmann et al., 2005). Criteria for transferring the product name were set to minimum 50% identity and query coverage of at least 70%. When the criteria were not met the gene product was assigned as hypothetical protein. Blast2Go 3.0.10 (Conesa et al., 2005) was used to perform InterproScan (Jones et al., 2014) and assigning GO-annotation. Carbohydrate Active enZymes (CAZymes) were predicted using the web-interface of dbCAN (Yin et al., 2012), which is based on the family classification in the CAZY database (http://www.cazy.org (Lombard et al., 2014)).

The annotated genome assembly has been submitted to NCBI/Genbank under the accession LFMY00000000 with the locus tag prefix UA08. It will be made public available when the present study is published.

5.2.6 Genome mining for SM clusters and putative PKS orthologue finding

Secondary metabolite gene clusters, and thereby the SM "backbone enzyme" genes encoding terpene synthases (TP), non-ribosomal peptide synthases (NRPS) and polyketide synthases (PKS), were identified using SMURF (Khaldi et al., 2010) and antiSMASH 2.0 (Blin et al., 2013). For the identification of PKSs additionally blastp analysis against a set of 543 PKS protein sequences (Delgado et al., 2012) were performed. The result was combined with the output of SMURF and antiSMASH. Putative PKS genes were additionally identified in four previously released *Talaromyces* genomes. Domain structures of identified PKSs were identified using the NCBI Conserved Domain Database (Marchler-Bauer et al., 2011).

For phylogenetic relationship between the *T. atroroseus* PKSs the individual KS domains were first aligned with MAFFT v6.902b (Katoh et al., 2002; Katoh & Toh, 2008) using the L-

INS-I option. The aligned KS domains were used for Bayesian phylogenetic analysis conducted in MrBayes v3.2.1 (Ronquist et al., 2012). Two independent Markov chain Monte Carlo runs, under the GTR model, were performed employing 3 chains each (2 heated and 1 cold). The chains were run for 50,000 generations sampling a total of 1,001 trees. Convergence was assessed in Tracer (Rambaut et al., 2014) and 50% majority consensus tree was generated combining the two runs, after the first 25% of the trees had been discarded as burnin.

Identification of orthologous PKS genes between the *Talaromyces* genera were identified using the Shuffle-LAGAN global chaining algorithm (Brudno et al., 2003) as described elsewhere (Hansen et al., 2012).

5.2.7 Synteny plots of azaphilone gene clusters

Microsynteny of the azaphilone clusters were detected by the Shuffle-LAGAN algorithm (Brudno et al., 2003) using the putative azaphilone clusters including flanking regions. Easyfig (Sullivan et al., 2011) were used to create microsynteny plot employing blastn identities between individual genes. The microsynteny was following manually verified and slightly altered using the results from the Shuffle-LAGAN analysis.

5.2.8 USER vector construction

All primers used in this study are listed in Table 5.1. All PCR amplifications was performed with PfuX7 polymerase (Nørholm, 2010). Plasmids were constructed with USER cloning (Nour-Eldin et al., 2006; Hansen et al., 2011) or USER fusion (Geu-Flores et al., 2007). Constructed plasmids follow the nomenclature described by Hansen *et al.* (Hansen et al., 2011) with some small updates (see Appendix C.1). The hygromycin resistance gene (*hph*) with p*TrpC* and t*TrpC* from *Aspergillus nidulans* as promotor and terminator was amplified from pRF-HU (Frandsen et al., 2008). It was amplified in two fragments using primer pairs KBR009+KBR015 and KBR010+KBR014 thereby introducing a point mutation removing the AsiSI restriction site in the middle of the *hph* gene. The two fragments were cloned into pU0002 (Hansen et al., 2011) using the PacI USER cassettes to generate pU4004, with regenerated PacI USER cassettes. Recombination sites respectively upstream of UA08_09244/*monA* and downstream of UA08_09245/*monB* were amplified with KBR048+KBR049 and KBR050+KBR051 from *T. atroroseus* gDNA in order to be used for

Table 5.1. List of primers used in this study.

Primer #	Primer name	Sequence, 5'-3'
KBR001	TA.mitA.part1.fw	AGAGCGAUGCCTACATTAAGACAAGAGGG
KBR002	TA.mitA.part1.rv	ATTCCAUGACCAGTTATCGGG
KBR003	TA.mitA.part2.fw	ATGGAAUATCATCAGTGAGGG
KBR004	TA.mitA.part2.rv	TCTGCGAUTGGCAGTCAAGACCTTCTCGC
KBR009	hph_v1-fw	GGGTTTAAUGGGTTTAAUTAAGTCCTCAGCGAATGCGTG CGATCGCGTGCATTCGGCCTCGACAGAAGATGATA
KBR010	hph_v1-rv	GGTCTTAAUTAATGCCTCAGCTACCCACGCCGAAACAAGC
KBR014	hph-PM-fw	ATGCGATUGCTGCGGCCGATCTTAGC
KBR015	hph-PM-rv	AATCGCAUCCATGGCCTCCGCGAC
KBR016	TA.mitA.UP-fw	GGGTTTAAUGGGCCGAGCAGAGAAAAGA
KBR017	TA.mitA.UP-rv	GGACTTAAUGCGGAGGAAGTTGTGGAAAG
KBR018	TA.mitA.DW-fw	GGCATTAAUACGATTTTAACGGGGGCAG
KBR019	TA.mitA.DW-rv	GGTCTTAAUACCCCAACAACCCAACCCA
KBR027	TA.mitA-T3	ACTGCATCCACTCCCTGTCCT
KBR028	TA.mitA-T4	TGCTCTCGTCGTCCTCCGTT
KBR036	mitA.3'-Probe-fw	TGAGGAGGGGAGCACGAGATTGA
KBR037	mitA.3'-Probe-rv	AAGAAGGCAGACGCTGAAATGGGA
KBR038	mitA.Probe-fw	TTGGACCGCGCATTGGTGTTT
KBR039	mitA.Probe-rv	GGCTCGCCATGTTTGTGATGGT
KBR048	TA.monA.UP-fw	GGGTTTAAUTTCTTTTTGGGGCCATCTT
KBR049	TA.monA.UP-rv	GGACTTAAUTTTGCACCTACTCATCGATT
KBR050	TA.monB.DW-fw	GGCATTAAUCCAATCAACACTCACGGCA
KBR051	TA.monB.DW-rv	GGTCTTAAUCCTATTCCCTACTCAACGCT
KBR070	TA.monA-T3	GGTGAGGAAGACATGGCGGA
KBR071	TA.monB-T4	GGTGGCGATAGGCGGAGTAA
KBR084	pDHX3.backbone.fw	AGGTAAAUCCCCACTACCGCATTAAG
KBR085	pDHX3.backbone.fw	AGCTGCTUCGTCGATTAAACCCTCAGCG
KBR086	pDHX3.AMA1.fw	AAGCAGCUGACGGCCAGTGCCAAGCT
KBR087	pDHX3.AMA1.rv	ATTGGGGUACTAACATAGCCATCAAATGCC
KBR088	pDHX3.AMA2.fw	ACCCCAAUGGAAACGGTGAGAGTCCAGTG
KBR089	pDHX3.AMA2.rv	AGTGTCUGCGATCGCTCTCACTGCGCGGT
KBR090	pDHX3.pTrpC.fw	AGACACUGCTAGTGGAGGTCAACACATC
KBR091	pDHX3.pTrpC.rv	ATTTGGAUGCTTGGGTAGAATAGGTAAGT
KBR092	pAN8-1.ble.fw	ATCCAAAUGGCCAAGTTGACCAGTGC
KBR093	pAN8-1.ble.rv	ATTTACCUCTAAACAAGTGTACCTGTGC
KBR107	Hyg588U	AGCTGCGCCGATGGTTTCTACAA
KBR108	Hyg588L	GCGCGTCTGCTGCTCCATACA
KBR150	Ble.ch-fw	AGTTGACCAGTGCCGTTCCG
KBR151	Ble.ch-rv	CGTAGGTCTCTTGACGACCGTTG
KBR216	TA.mitA-GapCH-fw	TCGCTTTCCACAACTTCCT
KBR217	TA.mitA-GapCH-rv	TCAACAGAAGAACTGCCCC

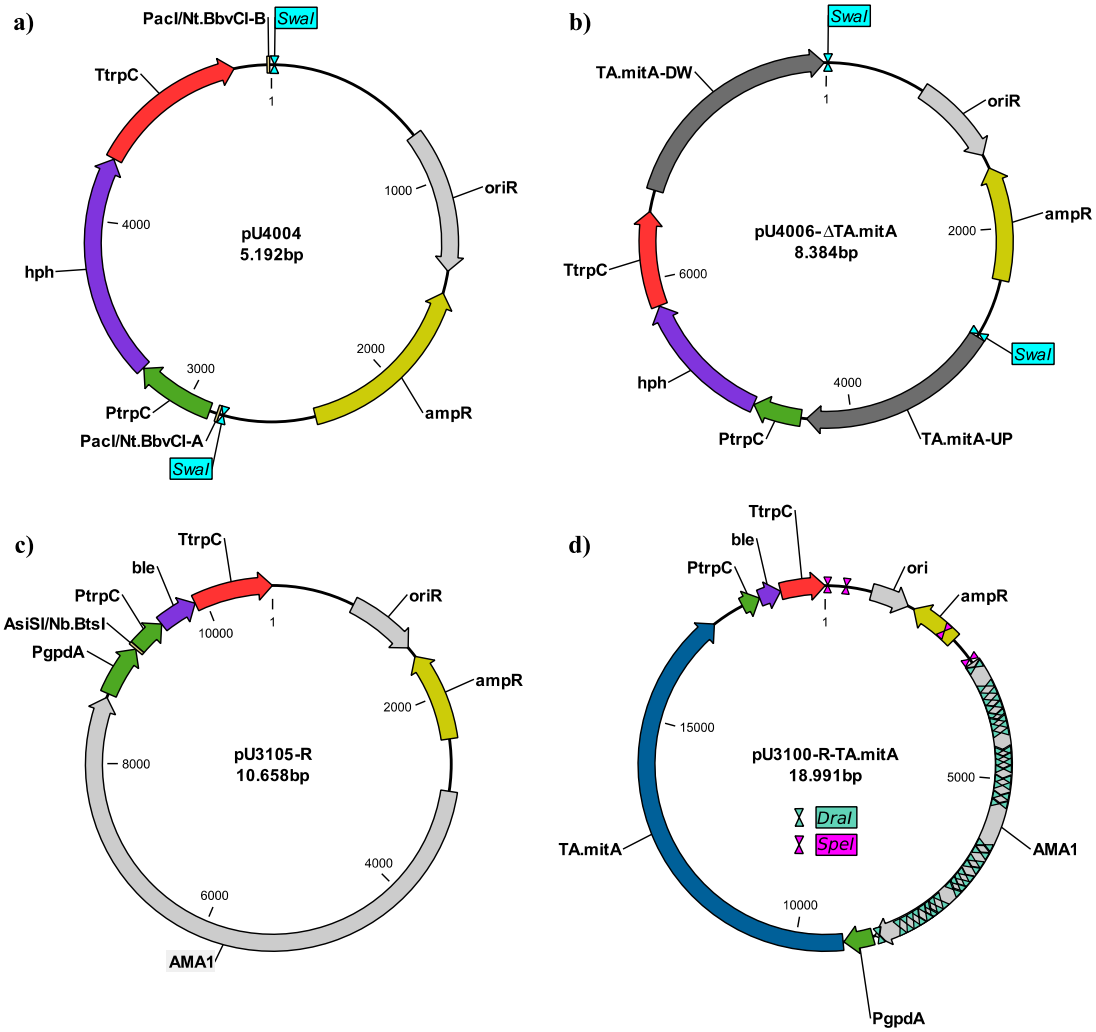


Figure 5.1. Vectors created and used in this study. USER cassettes are specified by the respective restriction and nicking enzyme combination needed. Restriction sites used in liberation of linear transformation fragments have been highlighted.

knockout of the putative dual FAS-complex genes *monA* and *monB*. Up- and downstream recombination sites of UA08_05984/*mitA* were amplified using KBR016+017 and KBR018+019. The PCR fragments were USER cloned into pU4004 using the PacI USER cassettes creating the plasmids pU4006-ΔTA.*monA*+*monB* and pU4006-ΔTA.*mitA*. The plasmids were linearized with SwaI to liberate the gene targeting substrates.

For reconstitution of MitA activity in *T. atroroseus*, an AMA1-based (Gems et al., 1991) plasmid containing the *ble* gene providing resistance towards phleomycin was created. Four fragments was amplified from pDHX3 (Holm, 2013) using primer pairs KBR084+085, KBR086+087, KBR088+089 and KBR090+091 and the *ble* resistance gene was amplified from pAN8-1 (Punt et al., 1988) using primers KBR092+093. The five fragments were

assembled with USER fusion (Geu-Flores et al., 2007) creating the plasmid pU3105-R (Figure 5.1c). The *mitA* gene was amplified from *T. atroroseus* gDNA in two fragments including its native terminator with primers KBR001-KBR004 and cloned into pU3105-R. The resulting pU3100-R-TA.*mitA* (Figure 5.1d) was transformed into the *mitA* deletant TAT3, both uncut as well as linearized with DraI and SpeI.

5.2.9 Strain construction

Protoplast generation and transformation was performed as described previously for *Aspergillus nidulans* (Tilburn et al., 1983; Nielsen et al., 2006). For transformations 100 µl protoplasts were mixed with 2-5 µg linear DNA and incubated on ice. 40% PEG in 1M Sorbitol, 50 mM Tris, 10 mM $CaCl_2$, pH 7.5 was added in two rounds; 100 µl followed by 15 min incubation on ice and 500 µl followed by 15 min incubation at RT. The protoplast-DNA mixture was added to ~6 ml soft (0.8% agar) sCYA (Czapek-Dox Yeast agar supplemented with 1 M Sorbitol) overlaid on sCYA plates (2% agar) and incubated at 30 °C. The day after the transformation the plates were overlaid with ~8 ml soft sCYA containing 300 µg/ml Hygromycin B (Invivogen, Toulouse, France) or 50 µg/ml Phleomycin (Invivogen, Toulouse, France) depending on required selection. Plates were incubated at 30 °C until transformants broke through the overlay. Transformants were transferred to fresh CYA plates containing the same antibiotic concentration. Diagnostic PCR was performed directly on the fungal mycelia to verify correct transformants (Nielsen et al., 2008). To further confirm correct and clean integration of DNA substrates southern blot was performed. Specifically genomic DNA (1-2 µg per lane) were digested to completion with NdeI and separated in 0.8% agarose gels by gel electrophoresis. The DNA were transferred to a Hybond+ membrane using a Vacu-Blot transfer unit (Biometra, Göttingen, Germany) according to manufacturer's protocol. The DNA was probed with three different probes, see Figure 5.7 for design. Probe 1 and probe 2 was amplified from *T. atroroseus* genomic DNA using primer KBR037+038 and KBR038+039.

Table 5.2. Strains constructed and used in this study.

Strain	Genotype	Source
IBT11181 / TAT1	Wild type	IBT Culture Collection
TAT2	*monA*Δ-*monB*Δ::P*trpC*::*hph*::T*trpC*	This study
TAT3	*mitA*Δ::P*trpC*::*hph*::T*trpC*	This study
TAT7	*mitA*Δ::P*trpC*::*hph*::T*trpC*, pU3100-R-TA.*mitA*	This study
TAT8	*mitA*Δ::P*trpC*::*hph*::T*trpC*, NN::P*gpdA*::TA.*mitA*::*ble*	This study
TAT9	*mitA*Δ::P*trpC*::*hph*::T*trpC*, NN::P*gpdA*::TA.*mitA*::*ble*	This study

They were designed to probe against *mitA* and upstream of the *mitA* upstream recombination site respectively. Probe 3 was amplified from pRF-HU using primers KBR107+108 and designed to probe against the *hph* gene for verification of single and correct integration of the *hph* gene.

5.2.10 Chemical analysis of secondary metabolites

Secondary metabolites were extracted from 7 day old colonies using the micro-extraction procedure developed by Smedsgaard *et al.* (Smedsgaard, 1997). The extracts were analysed by UHPLC-HRMS as described in Chapter 3 & 4. Extracted ion chromatograms were used to evaluate production of azaphilone pigments in the different strains.

5.3 Results

5.3.1 *Talaromyces atroroseus* genome sequence and gene calling

The genus of *Talaromyces* primarily constitutes saprophytic fungi and encompasses important and interesting species such as the opportunistic human pathogen *T. marneffei*, industrially interesting pigment producing species *T. atroroseus* (Chapter 3 and Frisvad et al., 2013) as well as species with high production of cellulolytic enzymes i.e. *Talaromyces cellulolyticus* (Inoue et al., 2014). Several strains of *Talaromyces atroroseus* and closely related species are recognized as industrial potential cell factory for MP production as a mycotoxin-free alternative to *Monascus* species (Mapari et al., 2005; Mapari et al., 2009b; Frisvad et al., 2013). *T. atroroseus* IBT11181 in particular have shown to be a good strain for the further investigation of its potential for pigment production as highlighted in Chapter 4 as well as in previous studies at DTU Systems Biology (Mapari et al., 2009a; Mapari et al., 2009b). *T. atroroseus* IBT11181 originally isolated from red sweet bell pepper was selected for genome sequencing to serve as model-organism to more deeply investigate *T. atroroseus* as a suitable pigment producing cell factory. *Talaromyces atroroseus* IBT11181 was shot-gun sequenced using Illumina HiSeq 2000 on a 180 bp paired-end library and a 6 kbp mate-paired library both with reads of 2x100 bp. The estimated genome size, using kmer-count with a kmer of 25, was by ALLPATHS-LG estimated 33.19 kbp. This gives an average sequencing depth of 193x. The genome was assembled into 48 scaffolds with a total assembly size of 30.85 Mb corresponding to ~93% of the estimated genome size. The minimum number of scaffolds covering at least

half the assembly size, the N50 value, is 7 scaffolds and the length of the shortest of these scaffolds, the L50 value, is 1,577,401 bp. These assembly statistics are in the same range as the related *T. aculeatus* sequenced and assembled with a similar strategy (Table 1). The CEGMA pipeline (Parra et al., 2007) identified 238 of the 248 CEGs (95.97%) to be found as complete in the *T. atroroseus* assembly. Adding the partial hits with alignment length less than 70% identified four additional CEGs making the partial CEGs identified to 97.58%. The high percentage of complete CEGs indicates that the draft genome assembly is good with a high completeness and is hence valid to use for whole genome analysis.

The *ab initio* gene prediction using AUGUSTUS resulted in 8997 and 8913 predicted genes when trained using *A. nidulans* and *F. graminearum* respectively. Gene prediction using FGENESH resulted in slightly more predicted genes, i.e. 9479, 9444 and 9357 genes when trained with *A. nidulans*, *F. graminearum* and *P. funicolosum* respectively. Genemark-ES on the other hand predicted significantly more genes counting to a total of 10746. The individual *ab*

Table 5.3. Genome assembly and annotation statistics for newly sequenced *Talaromyces atroroseus* and other available *Talaromyces* genomes

Species	***T. atroroseus***	***T. stipitatus***	***T. aculeatus***	***T. marneffei***	***T. marneffei***
Strain	**IBT 11181**	**ATCC10500**	**ATCC10409**	**PM1[1]**	**ATCC18224**
Sequenced by[2]	DTU	JCVI	JGI	TAMU	JCVI
Sequencing technology	Illumina, PE + MP	Sanger	Illumina, PE + MP	Sanger, Illumina, PacBio	Sanger
Assembly algorithms	ALLPATHS-LG	Celera Assembler	ALLPATHS-LG	Celera Assembler	Celera Assembler
Assembly size (Mbp)	30.85	35.69	37.27	28.34	28.64
Sequencing depth	193x	8.09x	91.8x	150x	8.8x
Number of scaffolds	48	820	49	216	452
Scaffold N50[3]	7	4	5	16	4
Scaffold L50[4] (Mbp)	1.58	4.36	2.93	0.68	3.34
GC content, overall (%)	44.35	45.90	46.51	46.70	46.38
GC content, coding (%)	50.47	48.04	49.73	48.78	48.90
Number of genes	9520	12574	13793	10051	10138
Average protein length (aa)	537	480	472	523	504
Exons per gene, mean	3.42	3.30	3.16	2.47	3.16

[1]PM1 second assembly with accession JPOX00000000 (Yang et al., 2014)
[2]DTU: Technical University of Denmark, JGI: Joint Genome Institute, U.S. Department of Energy, TAMU: Texas A&M University, JCVI: J. Craig Venter Institute
[3]Scaffold N50: The minimum number of sequences making up 50% of the genome assembly
[4]Scaffold L50: Length of the scaffold for, which scaffolds of this length and longer, covers at least 50% of the sum of all scaffolds.

initio gene predictions were merged into a consensus *ab* initio gene prediction using EvidenceModeler resulting in a total of 9519 protein encoding genes serving as the final gene prediction. The GC content of the resulting coding part of the genome was a bit higher than in other genome sequenced *Talaromyces* (Table 5.3) and significantly higher than the overall GC content, which was also expected.

5.3.2 Functional annotations

Assigning putative functions to genes *in silico* can be of great help to the scientific community. One used and simple strategy to assign product names to genes is to perform blastx/blastp comparison to the NCBI nr database and transfer the product name of the best hit to the new annotation. However the usage of a non-curated database for this purpose leads to a proportion of the genes with wrongly assigned gene product names due to propagation of wrongly gene assignments. The purpose of using the nr database is expectably to provide as much possible information as possible to the research community; however in my opinion it is wise to be a bit conservative in assigning these product names in order not to mislead other researchers. So instead I did the blastp comparison against the curated UniProtKB/SwissProt database (Boeckmann et al., 2005) and only transferred product names when blast hits satisfied the cut-offs specified in *Materials and methods* (Section 5.2.5). With this approach 1611 (17%) of the putative proteins were assigned with putative product names. Furthermore at least one GO-annotation were assigned to 6155 (65 %) of the protein encoding genes. InterProScan assigned functional and/or structural domains to 8672 (91 %) of the protein encoding genes.

5.3.3 Carbohydrate active enzymes

Plant pathogenic and saprophytic filamentous fungi rely on the ability to break down complex plant material to release mono- and oligosaccharides, which they can readily utilize in their metabolism. To break down complex plant material the filamentous fungi produce and secrete a broad range of Carbohydrate Active enZymes (CAZymes). The identification of CAZymes in *T. atroroseus* revealed that the total number of CAZymes is in the same range as other genome sequenced fungi within *Eurotiales*, with the exception of the citrus pathogen *Penicillium digitatum* (See figure 5.2), which have significantly lower number of CAZymes

(Marcet-Houben et al., 2012). As in other fungi the richest CAZy-type in *T. atroroseus* is the glycoside hydrolases (GH) which encompasses a total of 233 GH-domains.

The potential of using *T. atroroseus* as industrial pigment producing organisms is linked to the economically feasibility of the pigment production since food grade pigments is a low value product. Exploring cheap substrates for growth and pigment production is in relation to this of great importance. Cheap and easy available substrates are typically in the form of plant biomass. The carbohydrate of plants is structured in lignocellulose, a complex network of lignin, hemicellulose and cellulose. For the breakdown of lignocellulose to simple mono- and oligosaccharides a broad range of CAZy-families is needed. The GH-families involved in degradation of plant cell wall and most abundant in fungi include GH2, GH3, GH5, GH27, GH31, GH35, GH43, GH74, and GH78 (Zhao et al., 2014). *T. atroroseus* have CAZymes with domains from all of these important families and thus have potential for efficient growth on and utilization of complex lignocellulytic substrates.

Marine biomass in the form of *Saccharina japonica* has previously shown promising results as substrate for pigment production in the close related *T. amestolkiae* GT11 (General et al., 2014). Further studies of pigment production by *T. atroroseus* using cheap substrates are needed to explore the full potential.

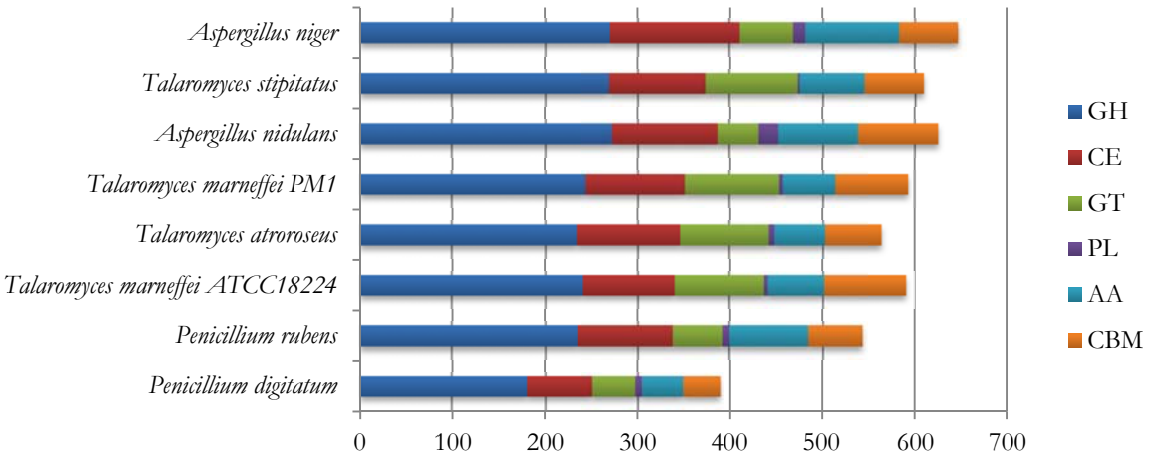


Figure 5.2. Distribution of CAZymes in *Talaromyces* and related species. Horizontal bars represent the number of identified CAZyme domains or modules. The organisms have been sorted in descending order in relation to the total number of CAZyme domains excluding the CBMs. GH: glycoside hydrolase, CE: carbohydrate esterase, GT: glycosyltransferase, PL: polysaccharide lyase, AA: auxiliary activity, CBM: carbohydrate binding module. Strain numbers not shown in figure: *Aspergillus niger* ATCC1015, *Aspergillus nidulans* FGSC A4. *Penicillium digitatum* PHI26, *Penicillium rubens* Wisconsin 54-1255, *Talaromyces stipitatus* ATCC10500.

5.3.4 Terpene synthases and non-ribosomal peptide synthases

Fungal secondary metabolites are of great interest to many researchers for their mycotoxic role in food and feed stock, the potential use as novel drug candidates or other beneficial uses. Secondary metabolites are usually grouped into three major classes; non-ribosomal peptides (NRPSs), terpenes and polyketides. Assisted by antiSMASH (Blin et al., 2013) and SMURF (Khaldi et al., 2010) key enzymes for secondary metabolism in *T. atroroseus* were identified. A total of seven putative terpene synthases were identified (Table 5.4). *T. atroroseus* produces large quantities of the compounds belonging to the group of sesquiterpene esters called purpurides (King et al., 1973; Wang et al., 2013) when grown on CYA and other standard mycological media. It is expected that the purpurides are produced by a biosynthetic cluster containing one of the identified terpene synthases. However, this potential linkage needs to be investigated further.

Table 5.4. Putative terpene synthases in *T. atroroseus*.
TIDS = Trans-Isoprenyl Diphosphate Synthase,
DMAT = Dimethylallyltransferase

Locus tag	Type
UA08_00320	Squalene cyclase
UA08_00830	Polyprenyl synthetase / TIDS
UA08_01226	Squalene synthase
UA08_01835	Terpene synthase
UA08_03139	Terpene synthase
UA08_07231	Polyprenyl synthetase / TIDS
UA08_08294	Tryptophan DMAT

A total of 11 NRPS genes were identified, see Figure 5.3. This is in the same range as found in other *Talaromyces* (*T. marneffei* ATCC18224: 11 NRPSs, *T. stipitatus* ATCC10500: 12 NRPSs), and a bit fewer as what is generally found in a number of genome sequenced Aspergilli, which contain between 11 and 20 putative NRPS genes (Khaldi et al., 2010). UA08_00560 only contains an adenylation and a condensation domain and it is likely that it is not a true functional NRPS. Two of the NRPSs contain five full modules i.e. UA08_06299 and UA08_07809, while UA08_03944 and UA08_06365 both contain four full modules. The UA08_08292 NRPS is clustered together with the DMAT encoded by UA08_08294 indicating a cluster capable of producing hybrid compounds of terpenoid / NRP nature.

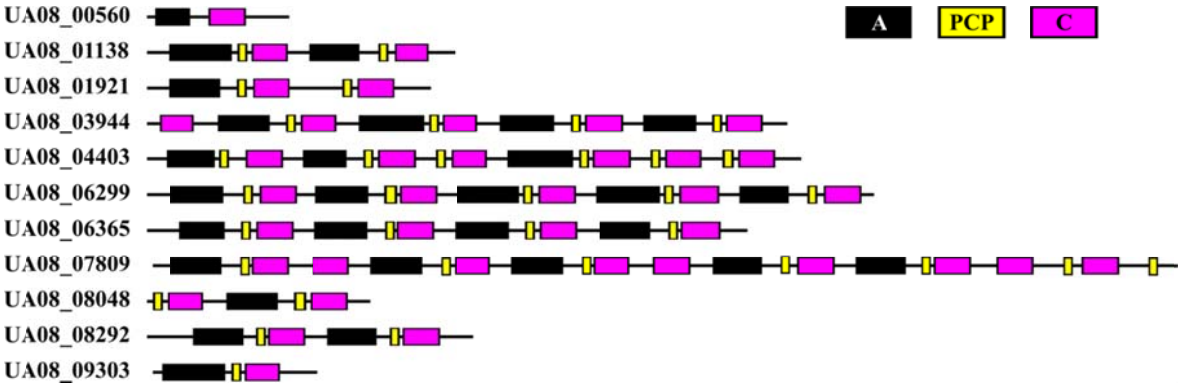


Figure 5.3. Domain architecture of *T. atroroseus* NRPSs. A: Adenylation domain, PCP: Peptidyl carrier protein, C: Condensation domain.

5.3.5 Polyketide synthases in genome sequenced *Talaromyces* species

The genomes of *T. stipitatus* ATCC10500, *T. marneffei* ATCC18224, *T. marneffei* PM1 *and T. aculeatus* ATCC10409 along with the newly sequenced *T. atroroseus* were mined for PKS encoding genes. The number of identified PKSs is summarized in Table 5.4. Detailed information on identified PKS genes along with domain structure of each PKS can be found in Figure 5.4 and Appendix C.2.

Table 5.5. Types and number of PKS encoding genes in *Talaromyces* genomes.

Species	*T. atroroseus*	*T. stipitatus*	*T. aculeatus*	*T. marneffei*	*T. marneffei*
Strain	IBT 11181	ATCC10500	ATCC10409	PM1	ATCC18224
Non-reducing PKSs	7	17	15	13	12
Reducing PKSs	7	19	20	11	11
PKS-NRPSs hybrids	5	3	1	2	2
Total PKSs	20	39	36	26	25

The PKS genes identified in *T. stipitatus* is in accordance with what has been found by others (Davison et al., 2012; Woo et al., 2010).

In the first published *Talaromyces marneffei* genome (Strain: PM1, Genbank Accession: AGCC01000000, (Woo et al., 2010)) Woo and co-workers identified 25 PKSs genes hereof two PKS/NRPS hybrids. In the recently released sequencing and assembly of *T. marneffei* PM1 (Genbank Accession: JPOX01000000, (Yang et al., 2014)) however two of the original identified PKS genes have not made the final assembly i.e. *pks4/alb1* which is responsible for melanin production (Woo et al., 2010) and *pks5* which has so far not been linked to metabolites. In addition to the 25 PKSs previously reported I identified the locus

GQ26_0500060 to cover two separate genes, a putative PKS encoding gene and a gene putatively encoding a RNA-directed DNA polymerase. In line with the numbering introduced by Woo et al (Woo et al., 2010) I name this PKS gene *pks26*. Comparing *pks26* to the first Sanger sequenced genome assembly of PM1 I found the *pks26* gene to be present but fragmented and spread over three different contigs (AGCC01001448, AGCC01001795, AGC01001492). This makes the number of PKS genes in the PM1 strain to a total of 26. The genome assembly of *T. marneffei* ATCC18224 contains orthologues for all of the original 25 *T. marneffei* PKS genes, but not the newly annotated *pks26*. This is a clear example showing that genome sequencing and annotation is error-prone and hence one should be careful to critically address the absence of specific gene or gene clusters before drawing conclusions.

The number of PKS genes in each *Talaromyces* species roughly reflects the genome assembly size; *T. stipitatus* and *T. aculeatus* being the two species with the largest genome assembly size also contain significantly more PKS genes than the two *T. marneffei* isolates and *T. atroroseus*. *T. atroroseus* with its 20 PKS genes contain less PKS genes than any of the other *Talaromyces* species but interestingly more of these being of the PKS-NRPS hybrid type.

5.3.6 Phylogeny of *Talaromyces atroroseus* PKSs

The KS domain is commonly used as the best domain for phylogenetic analysis of PKSs (Kroken et al., 2003; Baker et al., 2012) and was selected for the construction of phylogenetic tree of all PKSs in *T. atroroseus*. The inferred PKS phylogeny using Bayesian inference can be found in Figure 5.4. For reference KS domains from the *Aspergillus nidulans* WA and the *Aspergillus terreus* LovB was included. The WA PKS is involved in the biosynthesis of the yellow polyketide pigment YWA1 which is a precursor of the green conidial pigment in *A. nidulans* (Mayorga & Timberlake, 1990) while the LovB PKS is involved in the production of lovastatin in *A. terreus* (Hendrickson et al., 1999).

All of the PKSs contain the three domains required for a minimal PKS (McDaniel et al., 1994), β-ketosynthase (KS), acyltransferase (AT) and the acyl carrier protein (ACP). The phylogenetic analysis separated the PKSs into clades matching the clades reported by Kroken *et al.* (Kroken et al., 2003). In total two clades with non-reducing PKS genes and three clades with reducing PKS genes were identified in *T. atroroseus*. For the non-reducing PKSs the starter acyltransferase domains (SAT) have been manually annotated since these are not predicted by the CDD database. Together with the SAT domains, all the NR-PKSs have the expected product template (PT) domain involved in regioselective cyclization during chain length

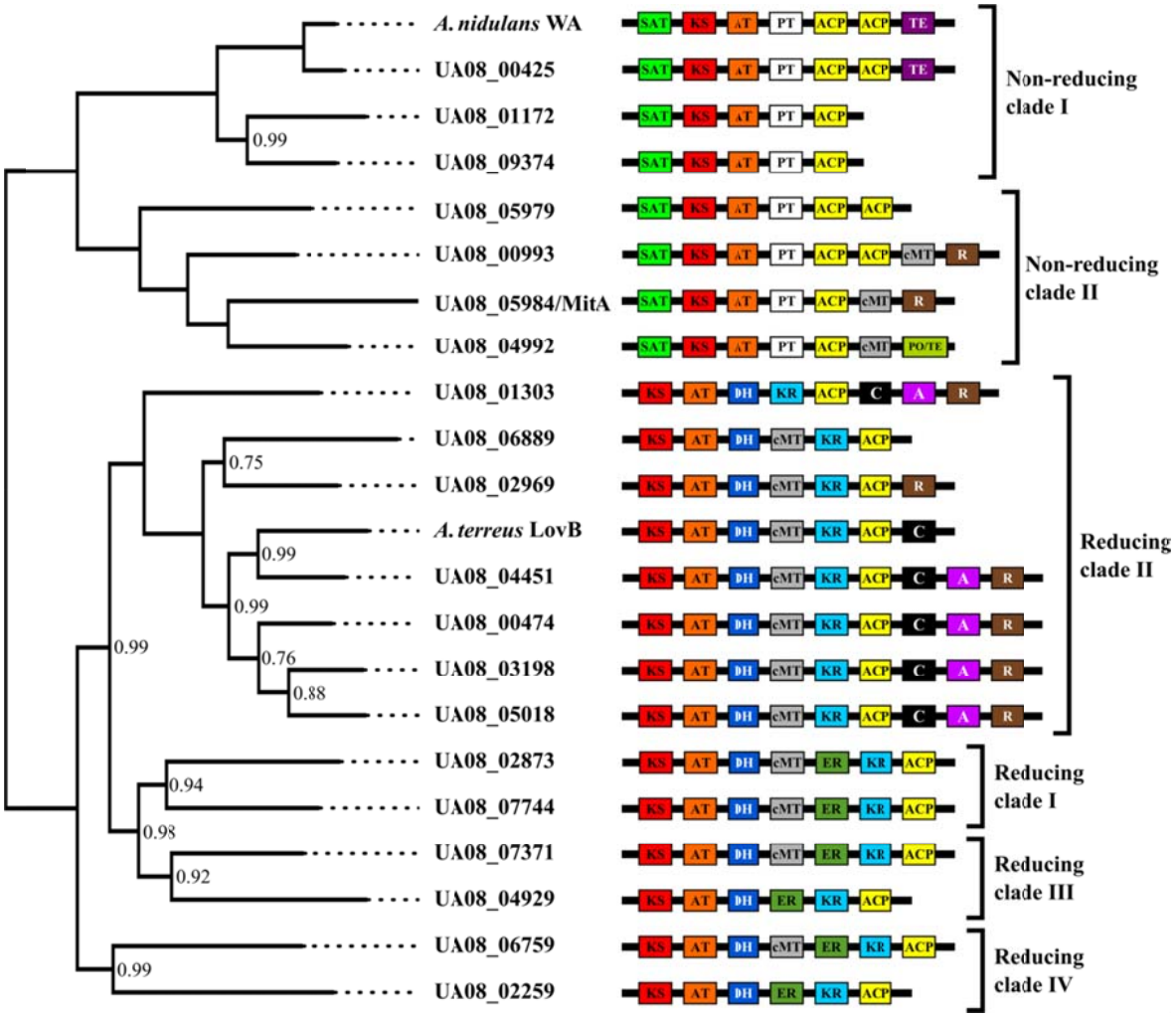


Figure 5.4. Phylogenetic relationship of KS domains from *T. atroroseus* PKS and PKS/NRPS hybrids. The tree was constructed with MrBayes as described in *Materials and methods* and midpoint rooted. Bayesian clade probabilities are shown on nodes when different from 1. The domain structure of each corresponding PKS is shown with the following abbreviations: SAT: starter acyl transferase, KS: β-ketosynthase, AT: acyl trasnferase, PT: Product template, ACP: Acyl carrier protein, DH: Dehydrogenase, ER: enoyl reductase, KR: ketoreductase, cMT: C-methyl transferase, C: condensation domain, A: adenylation domain, TE: Thioesterase, R: thioester reductase, PO: prolyl oligopeptidase.

elongation of aromatic polyketides (Crawford et al., 2009). Within the NR-PKS clade I UA08_00425 contain a releasing domain with thioesterase (TE) activity. It groups together with WA, which also uses a Claisen cyclase acting thioesterase domain as release mechanism (Fujii et al., 2001). Furthermore these contain tandem ACP domains. The other NR-PKS clade I PKSs have apparently lost their releasing domain and at the same time only contains one ACP domain.

The NR-PKS clade II contains two PKSs with a thioester reductase (R) domain. This domains is similar to the releasing domain typically found in NRPSs. UA08_04992 contains a

C-terminal esterase/lipase domain with partial hit to a prolyl oligopeptidase domain (pfam00326). Blast search of the UA08_04992 prolyl oligopeptidase domain resulted in best hits being the uncharacterized PKSs from *Capronia epimyces* (XP_007730098.1, ID: 65%), *Capronia coronate* (XP_007723599.1, ID: 63%), *Usnea longissimi* (AGI60156, ID: 55%) and *Zymoseptoria bravis* (KJX92105, ID: 49%). The *Aspergillus terreus* PKS produced from locus ATEG_03629 is also among the top hits (ID: 43 %). The ATEG_03629 encoded PKS was by Chiang *et al.* (Chiang et al., 2013) through heterologous expression in *A. nidulans* shown to produce the aromatic polyketide 5-methyl orsellinic acid. CDD domain prediction of the ATEG_03629 PKS also predicts the C-terminal to be of an esterase/lipase type with prolyl oligopeptidase activity. In line with the finding by Chiang *et al.* that the ATEG_03629 PKS can release its product, I propose that the C-terminal domain of UA08_04992 has thioesterase activity. The releasing domain might have slightly different mechanism than the thioesterases found in clade I of the NR-PKSs (Figure 5.4). Phylogenetic analysis of the putative release domain of UA08_04992 together with TE and R domains of the other *T. atroroseus* PKSs groups the UA08_04992 C-terminal domain closest together with the TE domains of UA08_00425 and WA (data not shown), which supports the proposed thioesterase activity. Furthermore it is noted that the C-methyl transferase (cMT) domain within the non-reducing PKSs only are present in the non-reducing PKS clade II.

Among the reducing PKSs from *T. atroroseus* all the PKS/NRPS hybrids group into the R-PKS clade II. LovB is also found to belong in R-PKS clade, which was also seen in the Kroken *et al.* study (Kroken et al., 2003). LovB, not being a PKS/NRPS hybrid, however still contains the condensation domain (C) of a NRPS but probably has lost the amino acid adenylation domain (A) as well as the reductive release domain (R). Apart from the PKS/NRPS hybrids two *T. atroroseus* PKSs without a functional NRPS part also groups within this clade. UA08_06889 has lost all the NRPS domains while UA08_02969 still contains the reduction domain. UA08_01303 being the only PKS/NRPS hybrid without a cMT domain is the first to split of in the subclade. The rest of the reducing PKSs contain all three reduction domains DH, ER and KR and groups together in reducing PKS clade I, III and IV.

5.3.7 Comparison of orthologues PKSs in *Talaromyces*

I compared the identified PKS genes from *T. atroroseus* with predicted PKS genes from other *Talaromyces* species. The orthology comparison revealed that the shared number of PKSs across the *Talaromyces* species is low. Similar observations has been done within other fungal

genera; e.g. *Fusarium* (Hansen et al., 2012; Brown et al., 2012), *Trichoderma* (Baker et al., 2012) and *Aspergillus* (Lin et al., 2012). Specifically, only three PKS genes are shared between all four *Talaromyces* species and each species contain a high fraction of unique PKSs between the compared species (Unique PKSs: *T. aculeatus,* 61%, *T. atroroseus,* 70%, *T. marneffei,* 38 % and *T. stipitatus,* 62 %). It is noteworthy that all five PKS-NRPS hybrids from *T. atroroseus* are unique to this species. These observations can and will probably change when more *Talaromyces* genomes are being sequenced and a more comprehensive analysis are done.

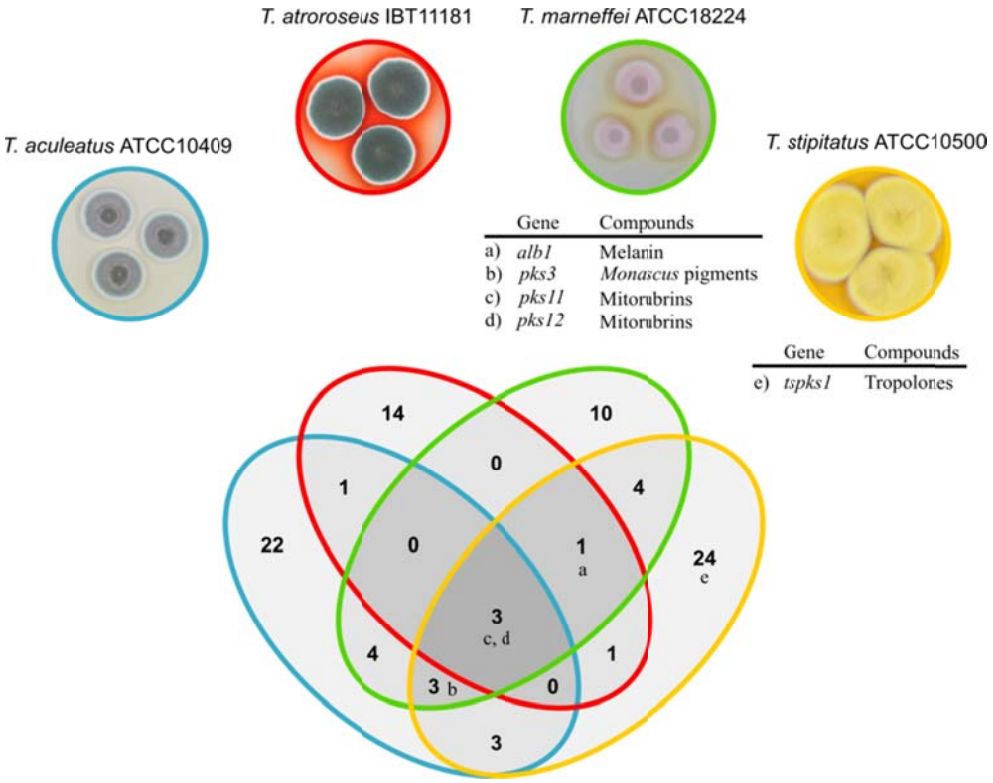


Figure 5.5. Venn diagram depicting orthologous PKSs identified in four *Talaromyces* genomes. PKS genes linked to their products is listed under the species where they were identified. For *Talaromyces marneffei* only strain ATCC18224 are shown in the figure. For orthology comparison of the PM1 strain, see Appendix C.2. Small letters in the Venn diagram indicates where the identified PKS genes belong.

5.3.8 *Talaromyces* PKSs linked to metabolites

A total of five of the PKS genes in *Talaromyces* has been linked to metabolites. These comprise the *T. marneffei* PM1 *pks4/alb1* gene (Woo et al., 2010), the *T. marneffei* PM1 *pks11* and *pks12* genes (Woo et al., 2012), the *T. marneffei* PM1 *pks3* gene (Woo et al., 2014) and the *T. stipitatus* ATCC18224 *tspks1/tropA* gene (Davison et al., 2012).

The *pks11* and *pks12* which are involved in biosynthesis of the mitorubrins are shared between all four *Talaromyces* species. A third PKS gene, the UA08_04929 ortholog, is shared among all species but no function has so far been assigned. Orthologs to the *alb1* gene, which is involved in the conidial pigmentation in the form of biosynthesis of melanin (Woo et al., 2010), are found in both *T. stipitatus* and *T. atroroseus* but is missing in *T. aculeatus*.

The *tspks1* is involved in biosynthesis of the tropolone stipitatic acid (Davison et al., 2012) and is only found in *T. stipitatus*. Orthologs of *pks3* are present in *T.stipitatus* and *T. aculeatus*, but not in *T. atroroseus*. This is striking since the *pks3* is linked to the production of the group of azaphilones comprising the red MPs (Woo et al., 2014) and *T. atroroseus* IBT11181 is known as an excellent producer of red *Monascus* and *Monascus*-like pigments (Mapari et al., 2009b; Chapter 3 and Frisvad et al., 2013). To verify that the gene calling or the PKS mining did not miss the MP PKS, a tblastx search against the *T. atroroseus* genome using *T. marneffei* PM1 *pks3* as query was performed. The closest related PKS gene in *T. atroroseus* showed out to be UA08_05984 with corresponding amino acid identity of 57%. However UA08_05984 is the *T. marneffei pks11* ortholog involved in biosynthesis of the mitorubrins as shown by Woo *et al.* (Woo et al., 2012). In conclusion no direct ortholog to the MP PKS is found in *T. atroroseus*.

5.3.9 The MP gene cluster in *Talaromyces* species is highly conserved

The inability to identify a PKS orthologous to *T. marneffei* PKS3 is stiking since *T. atroroseus* is a potent producer of MPs. This could be a symptom of an incomplete genome assembly not catching the MP cluster as seen for the *T. marneffei* PM1 assembly missing out on a PKS gene as described above. However, investigating the *T. atroroseus* genome I found the MP gene cluster intact with the exception of the missing PKS gene (Figure 5.6). The microsyntenic organisation of the MP gene cluster is well conserved within *Talaromyces* (Figure 5.6), with multiple rearrangements compared to the cluster organisation in *Monascus*. With the conserved microsynteny of the cluster within *Talaromyces* the MP PKS is expected to be found between UA08_09242 and UA08_09243. With PCR, using two set of primer pairs, I confirmed that the genomic distance between UA08_09242 and UA08_09243 is indeed around ~2.8 kb. With the very high degree of cluster conservation across the *Talaromyces* species I conclude that *T. atroroseus* seems to simply lack the MP PKS gene.

From the microsynteny comparison it is noted that all genes, previously proposed to be part of the MP gene cluster in *Monascus pilosus* (Balakrishnan et al., 2013; Balakrishnan et al., 2014a), are present in all *Talaromyces* species with the exception of *mppG*. The *mppG* encodes a

putative

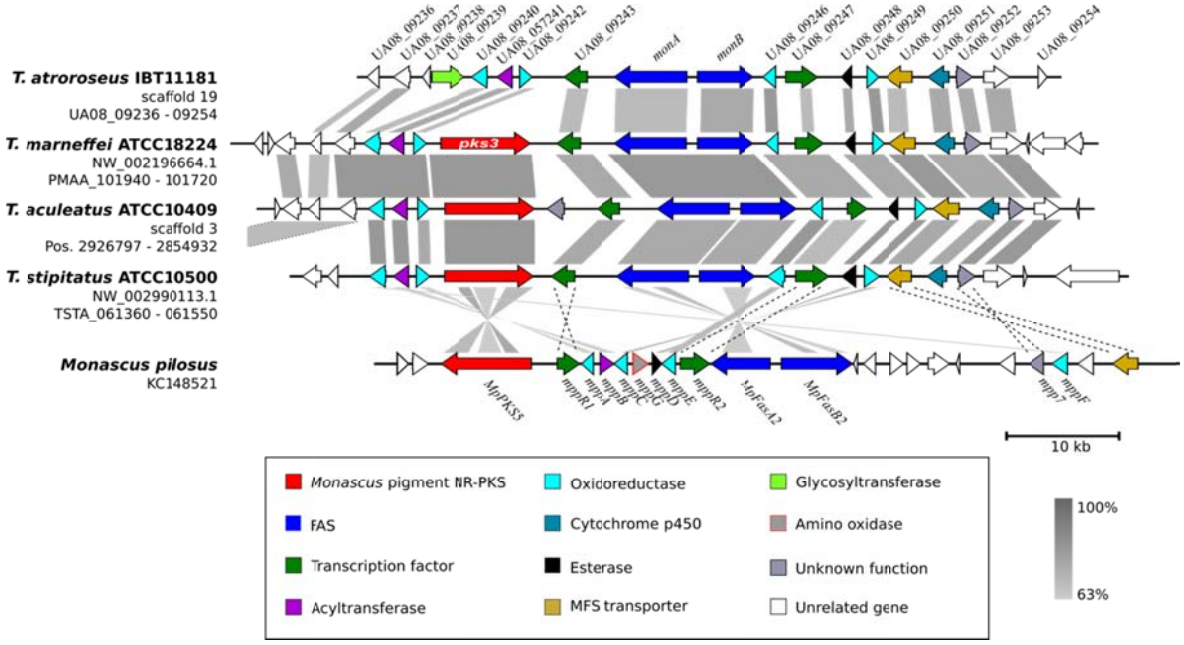


Figure 5.6. Microsynteny of azaphilone gene clusters across *Talaromyces* and *Monascus*. Grey tones specify the blastn identity. Dashed lines are used to link orthologous genes with blastn identity below the Easyfig treshold. TSTA_061420 and TSTA_061430 have been merged into one gene encoding the *MppR1* orthologue transcription factor in *T. stipitatus*.

amino oxidase, but the role in the MP biosynthesis is not clear. It is furthermore noted that *T. atroroseus* has an extra gene in the syntenic region of the MP cluster compared to the other *Talaromyces* species, i.e. UA08_09329. UA08_09329 contains a glycosyl transferase-like domain (Domain accession: cd03784). This is similar to domains used in glycosylation of the vancomycin group of antibiotics. N-glycosylated version of MPs has previously been isolated from *Monascus ruber* ATCC 96218 (Hajjaj et al., 1997), but it is not clear if UA08_09329 could be involved in a similar glycosylation process in *T. atroroseus*.

5.3.10 Knockout of UA08_09244 and UA08_09245 results in loss of red pigment production

The observation of the missing *Monascus* PKS in *Talaromyces atroroseus* IBT11181 made me speculate whether the putative MP cluster actually is responsible for the production of the red MPs in *T. atroroseus* IBT11181. To address this I simultaneously knocked-out the genes UA08_09244 and UA08_09245, encoding the dimeric fatty acid synthases α and β subunits. The knockout was verified by diagnostic PCR and southern blot (Figure 5.7) showing that the

hygromycin marker had replaced UA08_09244 and UA08_09245. The resulting knockout mutant was assigned the name TAT2. When grown on CYA, TAT2 showed no production of red pigments (Figure 5.9) as clearly seen by visual inspection directly on the plates, as well as in UHPLC-DAD analysed micro-extractions using the DAD trace at 525 nm. This confirms that the diminished MP gene cluster is responsible for the MP production, as seen in both *M. purpureus* (Balakrishnan et al., 2013) and *T. marneffei* PM1 (Woo et al., 2014). With the confirmation I will refer to the cluster as the MP cluster and I named the UA08_09244 and UA08_09245 genes *monA* and *monB*.

5.3.11 Knockout and complementation of UA08_05984 encoded PKS

After the confirmation of the involvement of the MP cluster in the biosynthesis of MPs in *T. atroroseus*, I was encouraged to find an explanation for the production of MPs with apparently no PKS present in the gene cluster. As described above, the PKS, which closest resemble the *Monascus* PKS from *T. marneffei* is the mitorubrin PKS encoded by UA08_05984.

The previously studied MP PKS and mitorubrin PKS have the same domain architecture of SAT-KS-AT-ACP-(ACP)-cMT-R, with the exception that the MP PKS have two ACP domains. This is true for all the analysed *Talaromyces* species. The azaphilone structural core of the mitorubrins closely resembles the core of the MPs and I hypothesized that UA08_05984 encoded PKS was capable of producing a compound, which can be used in the biosynthesis of both mitorubrins and MPs.

To test this hypothesis I UA08_05984 was knocked-out. In the UA08_05984 knockout strain the correct integration of *hph* and the loss of UA08_05984 were confirmed with diagnostic PCR as well as southern blot (Figure 5.7). The confirmed knockout strain was named TAT3 and it was unable to produce any red pigments when grown on CYA (Figure 5.9). This is strong evidence that UA08_05984 is involved in the biosynthesis of *Monascus* pigments in *Talaromyces atroroseus* IBT11181. To strengthen the evidence, I did complementation study overexpressing UA08_05984 in TAT3 under the control of the *gpdA* promoter from *A. nidulans*. Complementation was done by expressing *mitA* both on an AMA1 based self-replicating plasmid as well as by random integration.

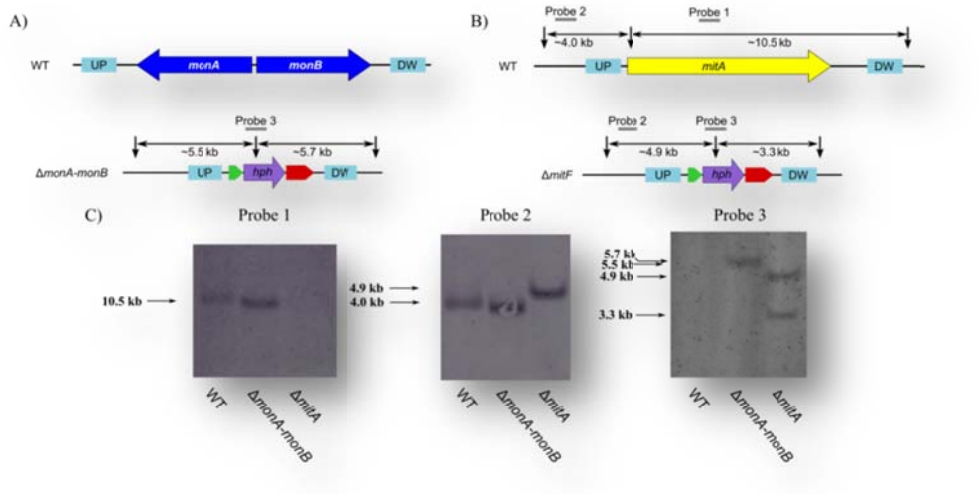


Figure 5.7. Southern blot confirmation of correct knockout of *monA + monB* and *mitA*. A) and B): Overview of the Southern blot design. Vertical arrows indicate NdeI restrictions sites and the length of the restriction fragments is indicated. C): Southern blots hybridized using three different probes. The figure is not drawn to scale.

The AMA1-based overexpression in TAT7 resulted in regain of the pigment production when grown under phleomycin selection. However, the pigment production was not at the same level as the wild type (Figure 5.9). This can be due to the instable nature of the AMA1-based plasmid given the used concentration of phleomycin, or that the *gpdA* promoter from *A. nidulans* is not optimal for constitutive transcription in *T. atroroseus*. When grown without selection only a faint shade of red pigments was seen at the centre of the colony (Figure 5.9) indicating that the AMA1-plasmid was lost frequently when not selected for.

Mutants with random integration and over-expression of UA08_05984 were verified by diagnostic PCR with primers amplifying *ble*. All mutants regained the ability for production of red MP. Two mutants named TAT8 and TAT9 were further examined for their red pigment production ability. Growing the complementation mutants on CYA resulted in production of red pigments that by visual inspection is in the same range as for the wild type (only TAT8 shown in Figure 5.9). Further inspections in the DAD trace showed that the pigment profile is very similar to that of the wild-type however with larger peaks indicating a higher production of red pigments. This apparent increase in pigment production could be a result of the site of integration. The increase was, however, observed for both TAT8 and TAT9. Thus, a more likely explanation of the increase of pigment production would be the placement of the constitutive strong *gpdA* promotor in front of *mitA*.

5.3.12 Chemical analysis of mutant strains

Micro-extractions from the mutant strains were analysed by UHPLC-DAD-HRMS and compared to the wild type pigment profile when grown on solid CYA. The wild type pigment profile was presented in Chapter 4 (Figure 4.3-4.4) but for the ease of readability the identified azaphilone compounds is repeated in Figure 5.8.

Extracted ion chromatograms (EIC) of the azaphilone compounds identified in Chapter 4 were used to verify presence or loss of individual pigments (Figure 5.10). The first observation is that production of compound **1** is only occurring in the wild type and the two complementation strains, TAT7 and TAT8 (Figure 5.9). I furthermore observed that both knockout strains TAT2 and TAT3 did not produce any of the known MPs; **4, 6** and **8.** These are however produced by the complementation strains TAT7 and TAT8. For compound **6** initial investigation of the EIC trace of m/z of 426.2262 shows that the compound is still produced in TAT2 and TAT3, but at lower amounts than in the wild type. However the EIC peaks in TAT2, TAT3 and TAT7 without selection are slightly shifted in retention time, eluting after 10.05 min. This is in contrast to the EIC peaks of the wild type, TAT7 with selection and TAT8 eluting after 10.11 min (see Appendix C.3). For TAT2, TAT3 and TAT7 without selection the EIC peaks turned out to come from the sodium adduct od another compound with $[M+H]^+$=404.2434 (see Appendix C.3). This is verified by the EIC of the sodium adduct of compound **6** (EIC: m/z = 448.2077, Figure 5.9), where no product is seen in TAT2, TAT3 or TAT7 without selection.

For the unknown MP compound **9** the EIC of m/z at 784.5864 shows that **9** is not linked to this mass since the red pigmentation is lost in TAT2, TAT3 and TAT7 without selection; a pattern that is not in concordance with the 784.5864 EIC-trace. Instead I link compound **9** to the EIC of m/z at 906.5185. This gives the proposed molecular formula $\mathbf{C_{58}H_{49}N_2O_5}$ for compound **9**. Unknown compound **5** and **7** show the same production pattern as the known MPs as well as **9.** It is thus confirmed that these belongs to the class of MPs and that both the *monA* and *monB* complex and *mitA* is required for the biosynthesis.

In TAT2 missing the dual FAS genes for generating the 3-oxoacyl moiety of the MPs it was expected that intermediates without the 3-oxoacyl moiety would accumulate. This was exactly what Balakrishnan *et al.* observed when doing the same knockout in *Monascus purpureus* (Balakrishnan et al., 2014b). The expected accumulation could either be in the form of **1, 2** or a similar compound. Surprisingly however, compound **2** was not produced in TAT2 and **1** was only produced in very scarce quantities. I was not able to identify any other accumulated

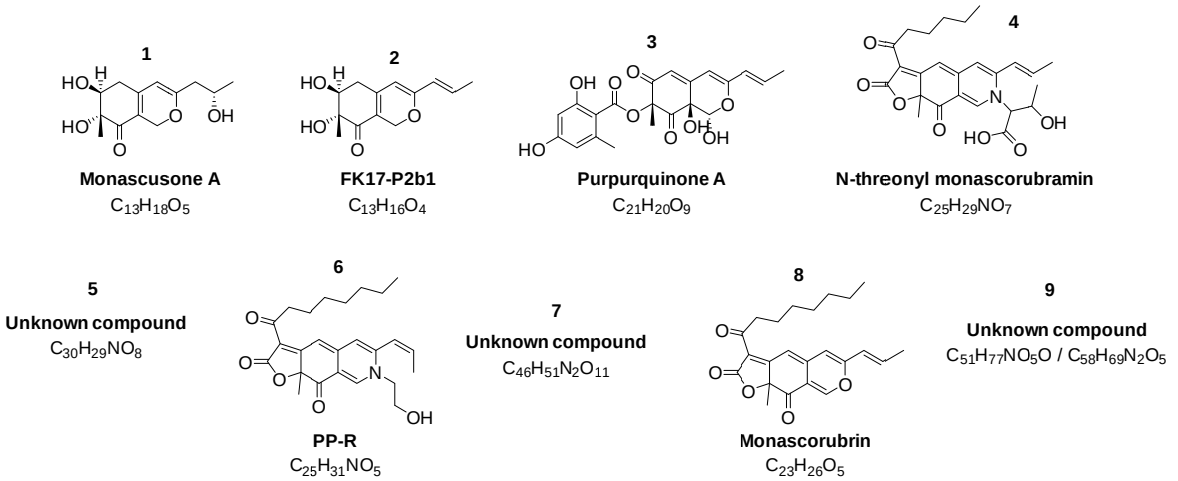


Figure 5.8. Azaphilone pigments produced by *Talaromyces atroroseus* when grown on CYA. Compounds identified in Chapter 4 as azaphilone pigments in micro-extractions of *T. atroroseus* after 7 days growth on CYA at 30 °C. For further reference and support for the pigment identification see Figure 4.4.

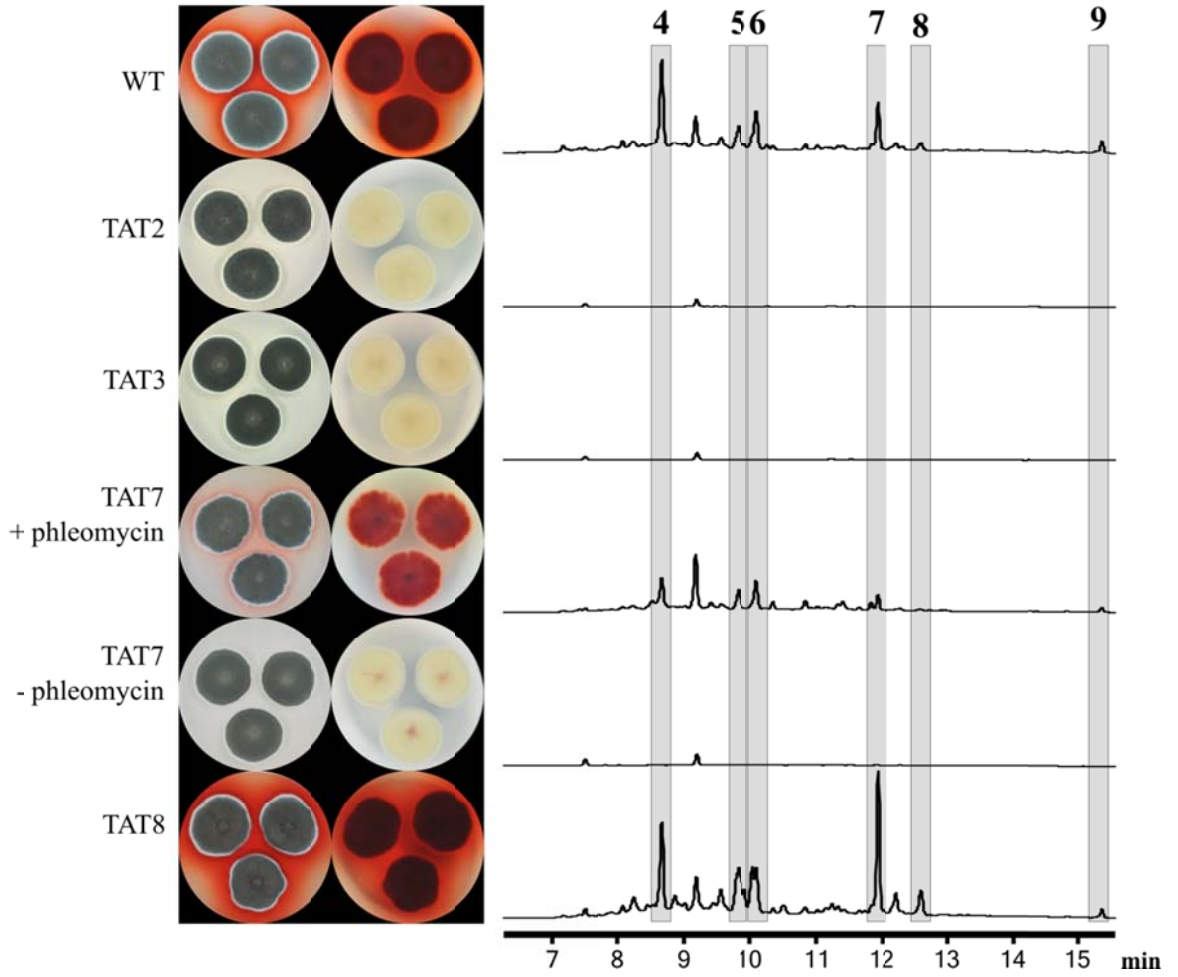
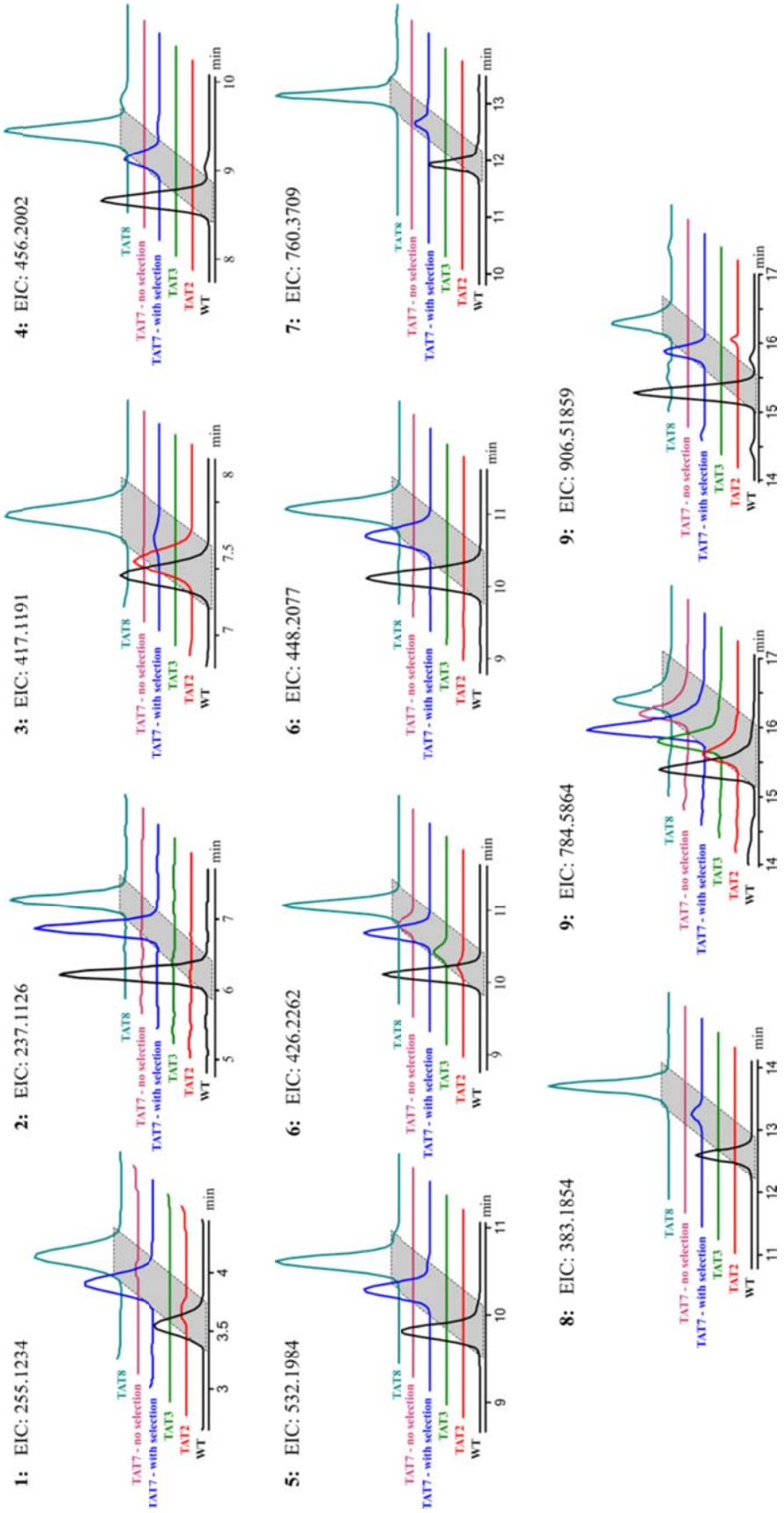


Figure 5.9. Production or red pigments in *Talaromyces atroroseus* wild type and mutants. Left: Pictures of cultures grown 7 days at 30 °C on CYA, obverse and reverse. Right: Extracted wave length chromatograms at 525 ± 4 nm showing the production and loss of red pigments. Numbers in bold are referring to azaphilone compound numbering introduced in Figure 5.7.


1: EIC: 255.1234
2: EIC: 237.1126
3: EIC: 417.1191
4: EIC: 456.2002
5: EIC: 532.1984
6: EIC: 426.2262
6: EIC: 448.2077
7: EIC: 760.3709
8: EIC: 383.1854
9: EIC: 784.5864
9: EIC: 906.51859
TAT8
TAT7 - no selection
TAT7 - with selection
TAT3
TAT2
WT
min

compounds corresponding to the bicylic azaphilone core. As **1** and **2** are also not being produced by TAT3 it could indicate that both *monA, monB* and *mitA* are needed for the biosynthesis of these. However the chemical formulas do not fit with addition of extra moieties to the azaphilone core and I therefore deem it a very unlikely explanation. An explanation for the apparently missing accumulation of azaphilone intermediates could be due to regulation. This regulation could either be in the form of rapidly convertion of the intermediate products into final azaphilones from the mitorubrin pathway, or down regulation of the *mitA* gene resulting in overall less azaphilone products being produced.

Looking at the production of Purpurquinone A (**3**), it being produced in the TAT2 strain. This was expected since it belongs to the mitorubrins and hence do not need the 3-oxoacyl moiety delivered by the *monA-monB* FAS complex. The production of compound **3** in TAT2 is comparable to the wild type. This indicates that there is no increased flux of metabolites through the mitorubrin biosynthetic pathway. No other azaphilone from the mitorubrinic class was identified.

5.3.13 Azaphilone production in *T. atroroseus*

Among the two classes of azaphilone compounds (the mitorubrins and the MPs) produced by *T. atroroseus* (Chapter 3, Chapter 4) the biosynthesis of MPs is the most studied, however, most of these studies focusing on the biosynthesis occurring in *Monascus* species (Balakrishnan et al., 2013; Xie et al., 2013; Liu et al., 2014; Balakrishnan et al., 2014a; Bijinu et al., 2014). In this study I have demonstrated that the biosynthesis of the two classes of azaphilones is entangled in *T. atroroseus*. More specifically I have shown that the mitorubrin PKS MitA is needed for the poruction of both MPs and mitorubrins. MitA likely produces a mitorubrin biosynthetic intermediate, which can be utilized in the production of MPs as shown in Figure 5.11.

Based on the highly conserved MP cluster across *Talaromyces* species, a likely explanation for the missing MP PKS in *T. atroroseus* is that is has been lost. The loss might have been driven by the redundancy of having two PKS genes in two separate gene clusters capable of delivering intermediates to both biosynthetic pathways. Natural selection have hence "thrown away" one of the PKS genes while retaining the PKS with the highest ecological advantage. The apparent redundant function of these large multifunctional enzymes is striking and the cross-chemistry between the two gene clusters of great interest to gain further understanding of the fungal potentials for generating the wide biodiversity of secondary metabolites.

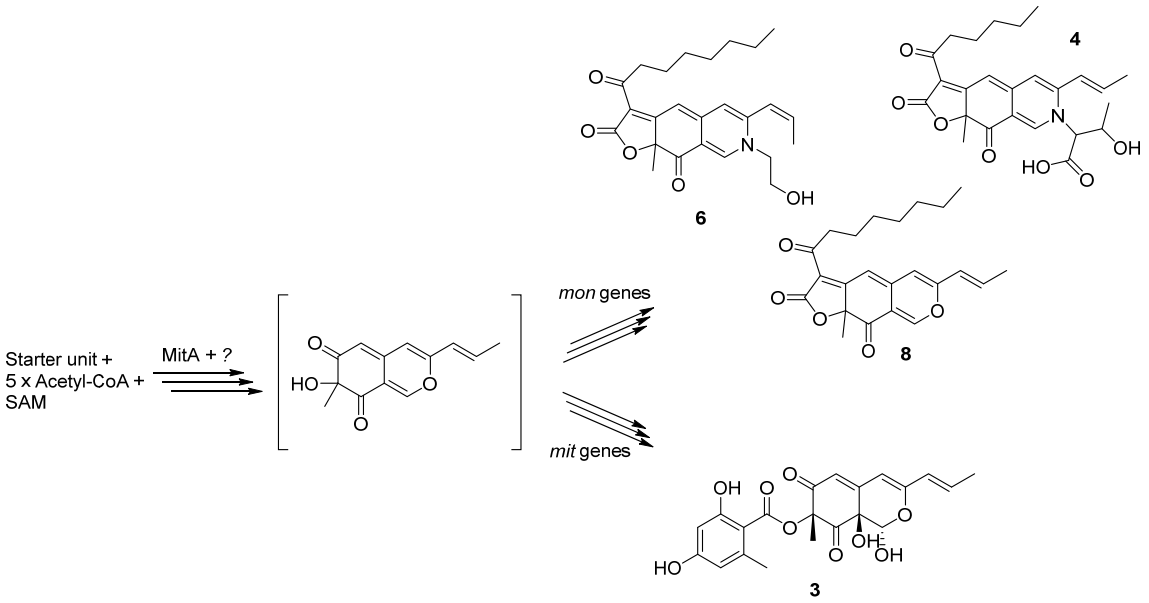


Figure 5.11. Simplified model of biosynthesis of azaphilones in *Talaromyces atroroseus* IBT11181. The NR-PKS encoded by *mitA* produces a compound that can be utilized by the biosynthetic pathway of both MP and mitorubrins. The shown intermediate is hypothetical and might also need other enzyme activities than MitA. Dedicated *mon* and *mit* gene products convert the common precursor to the observed pigments of the MP pathway and the mitorubrin pathway respectively. The compound in brackets is a hypothetical intermediate.

In the process of further understanding the biosynthesis of the azaphilones in *T. atroroseus* obvious next steps involves gene knockouts to identify biosynthetic intermediates. This might however prove more challenging and complex than studying the biosynthesis of MPs in a *Monascus* species, since the cross-chemistry between the MP and mitorubrin biosynthetic pathway may be more prevalent than at first glance. The potential cross-chemistry between the biosynthetic pathways coupled to the two azaphilone gene clusters can potentially make results from knockout studies hard to interpret.

Another strategy to further unravel the biosynthesis of azaphilones in *T. atroroseus* could be to heterologous express *mitA* to identify the product as the first intermediate in the biosynthetic pathways. Hashimoto *et al.* (Hashimoto et al., 2015) succeeded recently in heterologous expressing the mitorubrin PKS of *T. stipitatus* in *Aspergillus oryzae* generating four products. None of these compounds however folds up into the oxygenated bicyclic core of azaphilones. This is probably due to cross-chemical modifications by native enzymes of the host possibly combined with the lack of enzyme activities required for proper folding of the azaphilone core. In the same study the MP PKS of *T. stipitatus* however did not give any product. Such cross-chemical modifications are often observed when heterologous expressing SM clusters in other fungal hosts lowering the success-rate using this strategy. As MitA

contains a reductive putative release domain another strategy would be to produce and purify MitA and use it for *in vitro* studies of the enzyme activity and product formation.

In *T. marneffei* the knockout of MP PKS gene *pks3* results in the abolishment of production of MPs (Woo et al., 2014). Hence the mitorubrin PKS11 in *T. marneffei* is not capable of complementing the loss of PKS3. This suggests that the missing PKS in *T. atroroseus* has been preceded by adaptations to the MP biosynthetic genes so they can utilize an intermediate from the mitorubrin pathway. One interesting question to address for the future would be how widespread this adaptation and putatively loss of MP PKS is within *Talaromyces*. The rapid development of targeted capture for next generation sequencing provides possibilities for investigation cluster specific loss of the MP PKS in many *Talaromyces* species avoiding the need of whole genome sequencing.

5.4 Conclusion and perspectives

In the present study the key discovery is that *T. atroroseus* is missing the MP PKS, but is still a very capable producer of extracellular MPs. This is caused by an intriguing interplay between two SM biosynthetic pathways in *T. atroroseus*, i.e. the mitorubrin biosynthesis and the MP biosynthesis. The mitorubrin PKS, *mitA,* hence delivers polyketide scaffolds for the biosynthesis of both mitorubrins and MPs. To my knowledge this is the first example to date of a SM cluster, which has putatively lost its synthase but keeps producing the same range of compounds by "stealing" precursors from another biosynthetic pathway. Genes arranged in fungal secondary metabolite clusters is normally considered restricted to the biosynthesis of the compounds from that distinct cluster. The present finding add to the understanding of possible cross-chemistry between different SM biosynthetic clusters as well add important knowledge for the intelligent re-engineering of azaphilone pigment biosynthesis using combinatorial polyketide biosynthesis.

The way *mitA* complements the loss of MP PKS in *T. atroroseus* suggests that these two azaphilone PKSs are closely related. The evolutionary relationship between these PKS genes as well as the rest of the azaphilone clusters in *Talaromyces* will be further addressed in Chapter 6.

Acknowledgment

I am truly thankful to Associate Professor Simon Rasmussen for valuable discussion on the genome sequencing strategy and to Associate Professor Bent Petersen for running the ALLPATHS-LG genome assembly for me. Both are affiliated with the Metagenomics group at DTU Systems Biology.

5.5 References

Arai, T., Koganei, K., Umemura, S., Kojima, R., Kato, J., et al., **2013**. Importance of the ammonia assimilation by *Penicillium purpurogenum* in amino derivative *Monascus* pigment, PP-V, production. *AMB Express*, **3**: 19. doi:10.1186/2191-0855-3-19

Baker, S.E., Perrone, G., Richardson, N.M., Gallo, A. & Kubicek, C.P., **2012**. Phylogenomic analysis of polyketide synthase-encoding genes in *Trichoderma*. *Microbiology*, **158**: 147–154. doi:10.1099/mic.0.053462-0

Balakrishnan, B., Karki, S., Kim, S.C.H., Bora, J.S., Yoon, N.Y., et al., **2013**. Genetic localization and in vivo characterization of a *Monascus* azaphilone pigment biosynthetic gene cluster. *Applied Microbiology and Biotechnology*, **97**: 6337-6345. doi:10.1007/s00253-013-4745-9

Balakrishnan, B., Chen, C.-C., Pan, T.-M. & Kwon, H.-J., **2014a**. Mpp7 controls regioselective Knoevenagel condensation during the biosynthesis of *Monascus* azaphilone pigments. *Tetrahedron Letters*, **55**: 1640–1643. doi:10.1016/j.tetlet.2014.01.090

Balakrishnan, B., Kim, H.-J., Suh, J.-W., Chen, C.-C., Liu, K.-H., et al., **2014b**. *Monascus* azaphilone pigment biosynthesis employs a dedicated fatty acid synthase for short chain fatty acyl moieties. *Journal of the Korean Society for Applied Biological Chemistry*, **57**: 191–196. doi:10.1007/s13765-014-4017-0

Bijinu, B., Suh, J.-W., Park, S.-H. & Kwon, H.-J., **2014**. Delineating *Monascus* azaphilone pigment biosynthesis: oxidoreductive modifications determine the ring cyclization pattern in azaphilone biosynthesis. *RSC Advances*, **4**: 59405–59408. doi:10.1039/C4RA11713A

Blin, K., Medema, M.H., Kazempour, D., Fischbach, M. a, Breitling, R., et al., **2013**. AntiSMASH 2.0--a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Research*, **41**: W204–212. doi:10.1093/nar/gkt449

Boeckmann, B., Blatter, M.C., Famiglietti, L., Hinz, U., Lane, L., et al., **2005**. Protein variety and functional diversity: Swiss-Prot annotation in its biological context. *Comptes Rendus Biologies*, **328**: 882–899. doi:10.1016/j.crvi.2005.06.001

Brown, D.W., Butchko, R.A.E., Baker, S.E. & Proctor, R.H., **2012**. Phylogenomic and functional domain analysis of polyketide synthases in *Fusarium*. *Fungal Biology*, **116**: 318–331. doi:10.1016/j.funbio.2011.12.005

Brudno, M., Malde, S., Poliakov, A., Do, C.B., Couronne, O., et al., **2003**. Glocal alignment: finding rearrangements during alignment. *Bioinformatics*, **19**: i54–i62. doi:10.1093/bioinformatics/btg1005

Chiang, Y.M., Oakley, C.E., Ahuja, M., Entwistle, R., Schultz, A., et al., **2013**. An efficient system for heterologous expression of secondary metabolite genes in *Aspergillus nidulans*. *Journal of the American Chemical Society*, **135**: 7720–7731. doi:10.1021/ja401945a

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., et al., **2005**. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **21**: 3674–3676. doi:10.1093/bioinformatics/bti610

Crawford, J.M., Korman, T.P., Labonte, J.W., Vagstad, A.L., Hill, E.A., et al., **2009**. Structural basis for biosynthetic programming of fungal aromatic polyketide cyclization. *Nature*, **461**: 1139–1143. doi:10.1038/nature08475

Crawford, J.M., Thomas, P.M., Scheerer, J.R., Vagstad, A.L., Kelleher, N.L., et al., **2008**. Deconstruction of iterative multidomain polyketide synthase function. *Science (New York, N.Y.)*, **320**: 243–246. doi:10.1126/science.1154711

Davison, J., Cai, M., Song, Z., Yehia, S.Y. & Lazarus, C.M., **2012**. Genetic, molecular, and biochemical basis of fungal tropolone biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, **109**: 7642-7647. doi:10.1073/pnas.1201469109

Delgado, J.A., Al-Azzam, O., Denton, A.M., Markell, S.G. & Goswami, R.S., **2012**. A resource for the in silico identification of fungal polyketide synthases from predicted fungal proteomes. *Molecular Plant Pathology*, **13**: 494–507. doi:10.1111/J.1364-3703.2011.00760.X

Dufossé, L., Fouillaud, M., Caro, Y., Mapari, S.A. & Sutthiwong, N., **2014**. Filamentous fungi are large-scale producers of pigments and colorants for the food industry. *Current Opinion in Biotechnology*, **26**: 56–61. doi:10.1016/j.copbio.2013.09.007

Espinoza-Hernández, T.C., Rodriguez-Herrera, R., Aguilar-González, C.N., Lara-Victoriano, F., Reyes-Valdés, M.H., et al., **2013**. Characterization of three novel pigment-producing *Penicillium* strains isolated from the Mexican semi-desert. *African Journal of Biotechnology*, **12**: 3405–3413. doi:10.5897/AJB2013.12338

Frandsen, R.J.N., Andersson, J. a, Kristensen, M.B. & Giese, H., **2008**. Efficient four fragment cloning for the construction of vectors for targeted gene replacement in filamentous fungi. *BMC Molecular Biology*, **9**: 70. doi:10.1186/1471-2199-9-70

Frisvad, J.C., Yilmaz, N., Thrane, U., Rasmussen, K.B., Houbraken, J., et al., **2013**. *Talaromyces atroroseus*, a new species efficiently producing industrially relevant red pigments. *PLoS ONE*, **8**: e84102. doi:10.1371/journal.pone.0084102

Fujii, I., Watanabe, A., Sankawa, U. & Ebizuka, Y., **2001**. Identification of Claisen cyclase domain in fungal polyketide synthase WA, a naphthopyrone synthase of *Aspergillus nidulans*. *Chemistry and Biology*, **8**: 189–197. doi:10.1016/S1074-5521(00)90068-1

Fulton, T.M., Chunwongse, J. & Tanksley, S.D., **1995**. Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Molecular Biology Reporter*, **13**: 207–209. doi:10.1007/BF02670897

Gems, D., Johnstone, I.L. & Clutterbuck, a. J., **1991**. An autonomously replicating plasmid transforms *Aspergillus nidulans* at high frequency. *Gene*, **98**: 61–67. doi:10.1016/0378-1119(91)90104-J

General, T., Prasad, B., Kim, H., Vadakedath, N. & Cho, M., **2014**. *Saccharina japonica* , a potential feedstock for pigment production using submerged fermentation. *Biotechnology and Bioprocess Engineering*, **9**: 711–719. doi:10.1007/s12257-013-0709-2

Geu-Flores, F., Nour-Eldin, H.H., Nielsen, M.T. & Halkier, B. a., **2007**. USER fusion: A rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Research*, **35**: e55. doi:10.1093/nar/gkm106

Gnerre, S., Maccallum, I., Przybylski, D., Ribeiro, F.J., Burton, J.N., et al., **2011**. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proceedings of the National Academy of Sciences of the United States of America*, **108**: 1513–1518. doi:10.1073/pnas.1017351108

Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., et al., **2008**. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biology*, **9**: R7. doi:10.1186/gb-2008-9-1-r7

Hajjaj, H., Klaebe, A., Loret, M.O., Tzedakis, T., Goma, G., et al., **1997**. Production and identification of N-glucosylrubropunctamine and N-glucosylmonascorubramine from *Monascus ruber* and occurrence of electron donor-acceptor complexes in these red pigments. *Applied and Environmental Microbiology*, **63**: 2671–2678.

Hansen, B.G., Salomonsen, B., Nielsen, M.T., Nielsen, J.B., Hansen, N.B., et al., **2011**. Versatile enzyme expression and characterization system for *Aspergillus nidulans*, with the *Penicillium brevicompactum* polyketide synthase gene from the mycophenolic acid gene cluster as a test case. *Applied and Environmental Microbiology*, **77**: 3044–3051. doi:10.1128/AEM.01768-10

Hansen, F.T., Sørensen, J.L., Giese, H., Sondergaard, T.E. & Frandsen, R.J.N., **2012**. Quick guide to polyketide synthase and nonribosomal synthetase genes in *Fusarium*. *International Journal of Food Microbiology*, **155**: 128–136. doi:10.1016/j.ijfoodmicro.2012.01.018

Hashimoto, M., Wakana, D., Ueda, M., Kobayashi, D., Goda, Y., et al., 2015. Product identification of non-reducing polyketide synthases with C-terminus methyltransferase domain from *Talaromyces stipitatus* using *Aspergillus oryzae* heterologous expression. *Bioorganic and Medicinal Chemistry Letters*, **25**: 1381–1384. doi:10.1016/j.bmcl.2015.02.057

Hendrickson, L., Davis, C.R., Roach, C., Nguyen, D.K., Aldrich, T., et al., **1999**. Lovastatin biosynthesis in *Aspergillus terreus*: Characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene. *Chemistry and Biology*, **6**: 429–439. doi:10.1016/S1074-5521(99)80061-1

Holm, D.M.K., **2013**. *Development and implementation of novel genetic tools for investigation of fungal secondary metabolism*. PhD thesis. Technical University of Denmark.

Inoue, H., Decker, S.R., Taylor, L.E., Yano, S. & Sawayama, S., **2014**. Identification and characterization of core cellulolytic enzymes from *Talaromyces cellulolyticus* (formerly *Acremonium cellulolyticus*) critical for hydrolysis of lignocellulosic biomass. *Biotechnology for Biofuels*, **7**: 151. doi:10.1186/s13068-014-0151-5

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., et al., **2014**. InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, **30**: 1236–1240. doi:10.1093/bioinformatics/btu031

Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O., et al., **2005**. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research*, **110**: 462–467. doi:10.1159/000084979

Katoh, K., Misawa, K., Kuma, K. & Miyata, T., **2002**. MAFFT : a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, **30**: 3059–3066.

Katoh, K. & Toh, H., **2008**. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, **9**: 286–298. doi:10.1093/bib/bbn013

Khaldi, N., Seifuddin, F.T., Turner, G., Haft, D., Nierman, W.C., et al., **2010**. SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal genetics and biology*, **47**: 736–741. doi:10.1016/j.fgb.2010.06.003

King, T.J., Roberts, J.C. & Thompson, D.J., **1973**. Studies in mycological chemistry. Part XXX and last. Isolation and structure of purpuride, a metabolite of *Penicillium purpurogenum* Stoll. *Journal of the Chemical Society. Perkin Transactions 1*, **1**: 78–80. doi:10.1039/p19730000078

Kroken, S., Glass, N.L., Taylor, J.W., Yoder, O.C. & Turgeon, B.G., **2003**. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proceedings of the National Academy of Sciences of the United States of America*, **100**: 15670–15675. doi:10.1073/pnas.2532165100

Lin, S., Yoshimoto, M., Lyu, P.-C., Tang, C.-Y. & Arita, M., **2012**. Phylogenomic and domain analysis of iterative polyketide synthases in *Aspergillus* species. *Evolutionary Bioinformatics*, **8**: 373–387. doi:10.4137/EBO.S9796

Liu, Q., Xie, N., He, Y., Wang, L., Shao, Y., et al., **2014**. *MpigE*, a gene involved in pigment biosynthesis in *Monascus ruber* M7. *Applied Microbiology and Biotechnology*, **98**: 285–296. doi:10.1007/s00253-013-5289-8

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M. & Henrissat, B., **2014**. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, **42**: 490–495. doi:10.1093/nar/gkt1178

Mapari, S.A.S., Hansen, M.E., Meyer, A.S. & Thrane, U., **2008**. Computerized screening for novel producers of *Monascus*-like food pigments in *Penicillium* species. *Journal of Agricultural and Food Chemistry*, **56**: 9981–9989. doi:10.1021/jf801817q

Mapari, S.A.S., Meyer, A.S., Frisvad, J.C. & Thrane, U., **2009a**. Production of *Monascus*-like azaphilone pigments. Patent No. WO2009/026923 A2

Mapari, S.A.S., Meyer, A.S., Thrane, U. & Frisvad, J.C., **2009b**. Identification of potentially safe promising fungal cell factories for the production of polyketide natural food colorants using chemotaxonomic rationale. *Microbial Cell Factories*, **8**: 24. doi:10.1186/1475-2859-8-24

Mapari, S.A.S., Nielsen, K.F., Larsen, T.O., Frisvad, J.C., Meyer, A.S., et al., **2005**. Exploring fungal biodiversity for the production of water-soluble pigments as potential natural food colorants. *Current Opinion in Biotechnology*, **16**: 231–238. doi:10.1016/j.copbio.2005.03.004

Marcet-Houben, M., Ballester, A.-R., de la Fuente, B., Harries, E., Marcos, J.F., et al., **2012**. Genome sequence of the necrotrophic fungus *Penicillium digitatum*, the main postharvest pathogen of citrus. *BMC Genomics*, **13**: 646. doi:10.1186/1471-2164-13-646

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., et al., **2011**. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Research*, **39**: D225–D229. doi:10.1093/nar/gkq1189

Mayorga, M.E. & Timberlake, W.E., **1990**. Isolation and molecular characterization of the *Aspergillus nidulans* wA gene. *Genetics*, **126**: 73–79.

McDaniel, R., Ebert-khosla, S., Fu, H., Hopwoodt, D.A. & Khosla, C., **1994**. Engineered biosynthesis of novel polyketides: Influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. *Proceedings of the National Academy of Sciences of the United States of America*, **91**: 11542–11546.

Méndez, A., Pérez, C., Montañéz, J.C., Martínez, G., Aguilar, C.N., et al., **2011**. Red pigment production by *Penicillium purpurogenum* GH2 is influenced by pH and temperature. *Journal of Zhejiang University, Science B*, **12**: 961–968. doi:10.1631/jzus.B1100039

Mendez-Zavala, A., Contreras-Esquivel, J.C., Lara-Victoriano, F., Rodriguez-Herrera, R. & Aguilar, C.N., **2007**. Fungal production of the red pigment using a xerophilic strain *Penicillium purpurogenum* GH-2. *Revista Mexicana De Ingenieria Quimica*, **6**: 267–273.

Nielsen, J.B., Nielsen, M.L. & Mortensen, U.H., **2008**. Transient disruption of non-homologous end-joining facilitates targeted genome manipulations in the filamentous fungus *Aspergillus nidulans*. *Fungal Genetics and biology*, **45**: 165–170. doi:10.1016/j.fgb.2007.07.003

Nielsen, M.L., Albertsen, L., Lettier, G., Nielsen, J.B. & Mortensen, U.H., **2006**. Efficient PCR-based gene targeting with a recyclable marker for *Aspergillus nidulans*. *Fungal genetics and biology*, **43**: 54–64. doi:10.1016/j.fgb.2005.09.005

Nørholm, M.H.H., **2010**. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. *BMC Biotechnology*, **10**: 21. doi:10.1186/1472-6750-10-21

Nour-Eldin, H.H., Hansen, B.G., Nørholm, M.H.H., Jensen, J.K. & Halkier, B.A., **2006**. Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Research*, **34**: e122. doi:10.1093/nar/gkl635

Parra, G., Bradnam, K. & Korf, I., **2007**. CEGMA: A pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics*, **23**: 1061–1067. doi:10.1093/bioinformatics/btm071

Parra, G., Bradnam, K., Ning, Z., Keane, T. & Korf, I., **2009**. Assessing the gene space in draft genomes. *Nucleic Acids Research*, **37**: 289–297. doi:10.1093/nar/gkn916

Punt, P.J., Mattern, I.E. & Van Den Hondel, C.A.M.J.J., **1988**. pAN series from *gpdA* gene. *Fungal Genetics Newsletter*, **35**: 25–30.

Rambaut, A., Suchard, M., Xie, D. & Drummond, A.J., **2014**. Tracer v1.6. Available from http://beast.bio.ed.ac.uk/Tracer

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., et al., **2012**. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**: 539–542. doi:10.1093/sysbio/sys029

Salamov, A.A. & Solovyev, V. V., **2000**. Ab initio gene finding in *Drosophila* genomic DNA. *Genome Research*, **10**: 516–522. doi:10.1101/gr.10.4.516

Samson, R.A., Houbraken, J., Thrane, U., Frisvad, J.C. & Andersen, B., **2010**. *Food and Indoor Fungi*, CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands.

Santos-Ebinuma, V.C., Roberto, I.C., Teixeira, M.F.S. & Pessoa Jr., A., **2013a**. Improving of red colorants production by a new *Penicillium purpurogenum* strain in submerged culture and the effect of different parameters in their stability. *Biotechnology Progress*, **29**: 778–785.

Santos-Ebinuma, V.C., Teixeira, M.F.S. & Pessoa Jr., A., **2013b**. Submerged culture conditions for the production of alternative natural colorants by a new isolated *Penicillium purpurogenum* DPUA 1275. *Journal of Microbiology and Biotechnology*, **23**: 802–810.

Smedsgaard, J., **1997**. Micro-scale extraction procedure for standardized screening of fungal metabolite production in cultures. *Journal of Chromatography A*, **760**: 264–270.

Smit, A.F.A. & Hubley, R., **2014**. Repeatmodeler 1.0.8. [Open source software]. http://www.repeatmasker.org/RepeatModeler.html

Smit, A.F.A., Hubley, R. & Green, P., **2014**. RepeatMasker 4.0.5. [Open source software]. http://www.repeatmasker.org/

Stanke, M., Diekhans, M., Baertsch, R. & Haussler, D., **2008**. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics (Oxford, England)*, **24**: 637–644. doi:10.1093/bioinformatics/btn013

Stanke, M., Schöffmann, O., Morgenstern, B. & Waack, S., **2006**. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics*, **7**: 62. doi:10.1186/1471-2105-7-62

Sullivan, M.J., Petty, N.K. & Beatson, S. a, **2011**. Easyfig: a genome comparison visualizer. *Bioinformatics (Oxford, England)*, **27**: 1009–1010. doi:10.1093/bioinformatics/btr039

Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y.O. & Borodovsky, M., **2008**. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Research*, **18**: 1979–1990. doi:10.1101/gr.081612.108

Tilburn, J., Scazzoccbio, C., Taylor, G.G., Zabicky-zissman, J.H., Robin, A., et al., **1983**. Transformation by integration in *Aspergillus nidulans*. *Gene*, **26**: 205-221. doi: 10.1016/0378-1119(83)90191-9

Wang, H., Wang, Y., Liu, P., Wang, W., Fan, Y., et al., **2013**. Purpurides B and C, two new sesquiterpene esters from the aciduric fungus *Penicillium purpurogenum* JS03-21. *Chemistry and biodiversity*, **10**: 1185–1192. doi:10.1002/cbdv.201200175

Woo, P.C.Y., Lam, C.-W., Tam, E.W.T., Lee, K.-C., Yung, K.K.Y., et al., **2014**. The biosynthetic pathway for a thousand-year-old natural food colorant and citrinin in *Penicillium marneffei*. *Scientific Reports*, **4**: 6728. doi:10.1038/srep06728

Woo, P.C.Y., Lam, C.-W., Tam, E.W.T., Leung, C.K.F., Wong, S.S.Y., et al., **2012**. First discovery of two polyketide synthase genes for mitorubrinic acid and mitorubrinol yellow pigment biosynthesis and implications in virulence of *Penicillium marneffei*. *PLoS Neglected Tropical Diseases*, **6**: e1871. doi:10.1371/journal.pntd.0001871

Woo, P.C.Y., Tam, E.W.T., Chong, K.T.K., Cai, J.J., Tung, E.T.K., et al., **2010**. High diversity of polyketide synthase genes and the melanin biosynthesis gene cluster in *Penicillium marneffei*. *The FEBS Journal*, **277**: 3750–3758. doi:10.1111/j.1742-4658.2010.07776.x

Xie, N., Liu, Q. & Chen, F., **2013**. Deletion of *pig*R gene in *Monascus ruber* leads to loss of pigment production. *Biotechnology Letters*, **35**: 1425–1432. doi:10.1007/s10529-013-1219-1

Yang, E., Chow, W.-N., Wang, G., Woo, P.C.Y., Lau, S.K.P., et al., **2014**. Signature gene expression reveals novel clues to the molecular mechanisms of dimorphic transition in *Penicillium marneffei*. *PLoS Genetics*, **10**: e1004662. doi:10.1371/journal.pgen.1004662

Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., et al., **2012**. DbCAN: A web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Research*, **40**: W445-451. doi:10.1093/nar/gks479

Zhao, Z., Liu, H., Wang, C. & Xu, J., **2014**. Correction: comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics*, **15**: 6. doi:10.1186/1471-2164-15-6

Chapter 6

Evolution of azaphilone pigment gene clusters within *Eurotiales*

As presented in the previous chapter, the polyketide backbone of the *Monascus* pigments in *T. atroroseus,* is produced by the mitorubrin PKS, MitA. As briefly discussed the redundancy and cross-chemistry between the two azaphilone gene clusters in *T. atroroseus* is probably the driving force leading to the loss of the *Monascus* pigment PKS in *T. atroroseus*. In this chapter the evolutionary origin of the azaphilone clusters in *Talaromyces* and *Monascus* is addressed. This is done with offset in the evolutionary faith of the azaphilone non-reducing PKSs in *Eurotiales*. This chapter is together with chapter 5 in the process of being condensed into a research manuscript.

6.1 Introduction

Filamentous fungi are capable of producing a wide variety of secondary metabolites (SMs). The genetic machinery involved in the pathway to a particular SM is often clustered on the genome (Keller & Hohn, 1997). The difference in the distribution of such gene clusters in closely related species indicates that the clusters are subject to a high degree of duplication and loss, and/or lateral gene transfer. For instance species within genera such as *Talaromyces, Aspergillus* and *Fusarium* share remarkably few PKS genes with species from the same genera (Brown et al., 2012; Lin et al., 2012, Chapter 5).

Several *Talaromyces* species produce two distinct groups of pigmented azaphilones, mitorubrins and *Monascus* pigments (MP). In *T. marneffei* the two groups of azaphilones have been linked to two separate polyketide gene clusters (Woo et al., 2012; Woo et al., 2014). In Chapter 5 I found that the MP cluster in *T. atroroseus* is missing the PKS, which is responsible for the azaphilone scaffold in the MPs. Instead, the azaphilone scaffold in the MPs is created and delivered by the mitorubrin PKS, MitA. The MitA however also delivers azaphilone scaffolds for the production of the mitorubrin-type azaphilone, Purpurquinone A. This raises some interesting questions. Did *T. atroroseus* loose the MP PKS or was it acquired by the other *Talaromyces* species after the speciation? How is the evolutionary relationship of the NR-PKS from the mitorubrin pathway and the MP pathway, since the mitorubrin PKS can complement the missing MP PKS in *T. atroroseus*? Is this evolutionary relationship shared by more biosynthetic genes in the two azaphilone gene clusters? These are some of the questions I will address in this chapter.

The dual role of MitA in *T. atroroseus* strongly indicates an evolutionary coherence of the PKS genes in the two azaphilone clusters. The chemical group of MPs has so far only been reported produced by species from the *Monascus* genus as well as a subset of *Talaromyces* species formerly attributed to the *Penicillium* subgenus *Biverticillium*. Mitorubrins have long been known to be produced by many species belonging to the *Talaromyces* genus (Buechi et al., 1965; Frisvad et al., 1990), but is however also produced by members of the *Hypoxylon* genus including *Hypoxylon fragiforme* (Steglich et al., 1974). *Hypoxylon* belonging to the Xylariales under the *Sordariomycetes*, however, is only distant related to *Talaromyces* and since no *Hypoxylon* genome was public available at the initiation of this study it was decided to only focus on the evolution of the azaphilone gene clusters in *Eurotiales*.

Within *Eurotiales* no other genera than *Talaromyces* and *Monascus* have so far been linked to the production of either mitorubrins or MPs. The capability of producing compounds belonging to the class of azaphilones are however widely distributed across the fungal kingdom (Osmanova et al., 2010), and within *Eurotiales* other azaphilones has been linked to their biosynthetic clusters, i.e. azanigerones in *Aspergillus niger* (Zabala et al., 2012), asperfuranones in *Aspergillus nidulans* (Chiang et al., 2009) and citrinin in *Monascus purpureus* (Shimizu et al., 2005).

In this study the evolutionary coherence of the mitorubrin gene cluster and the MP gene cluster within *Eurotiales* is investigated. Phylogenetic trees of homologous gene pairs between the two azaphilone clusters were constructed by species tree-aware Bayesian inference, and the evolutionary history of the gene clusters were addressed by reconciliation of the gene trees to the corresponding species tree. Together with further support, these reconciliations were used to determine likely evolutionary scenario for the azaphilone clusters within *Eurotiales*.

6.2 Methods

6.2.1 Species tree and divergence-time estimation

Species phylogeny and divergence-time estimation was generated using the BEAST 2.1.3 (Bouckaert et al., 2014) software platform, which implement a Bayesian Markov chain Monte Carlo (MCMC) method. Multi gene approach using eight different genes were applied. Four of the selected genes are commonly used and well established in fungal species phylogeny, i.e. the genes encoding β-tubulin, RNA-polymerase large subunit, RNA polymerase second large subunit and elongation factor 1α. Additionally four genes (the orthologs of *Saccharomyces cerevisiae's* KOG1, VPS53, CLU1 and RFA1), identified by Capella-Gutiérrez *et al.* (Capella-Gutiérrez et al., 2014) as a robust set of phylogenetic markers in *Ascomycota,* were added to the analysis providing a robust set of eight genes for the multi locus tree construction. The individual genes were aligned with the RevTrans 2.0 (Wernersson & Pedersen, 2003) implementation of MAFFT (Katoh et al., 2002; Katoh & Toh, 2008). RevTrans aligns the genes at the amino acid level and then 'reverse translate' the alignment to the corresponding nucleotide alignment. By this approach the resulting nucleotide alignment is a codon position specific alignment.

For the divergence-time estimation secondary calibration points from work by Prieto and Wedin (Prieto & Wedin, 2013) were used. Specifically, *Saccharomyces cerevisiae, Magnaporthe oryzae, Fusarium graminearum, Cladophilalophora carrionii, Coccidioides posadasii* and *Trichophyton rubrum* were selected as representatives to specify the divergence dates priors obtained from the Prieto and Wedin study. The GTR+I+G model for molecular evolution of DNA were used, and analysis was run under the uncorrelated relaxed clock model with calibrated Yule model (Heled & Drummond, 2012) as tree prior. The MCMC was run for 50 million generations in 3 parallel runs with sampling every 25000 generation. Assessment of the convergence of the BEAST runs was performed with Tracer (Rambaut et al., 2014) and parallel runs were combined with LogCombiner discarding the first 20% of the trees as burn-in. The maximum clade credibility (MCC) tree was extracted and median node heights and HPD% intervals were summarized and assigned with TreeAnnotator.

6.2.2 Identification of putative paralogs between azaphilone clusters

Under the hypothesis that the two azaphilone SM clusters in *Talaromyces* evolved from the same ancestral SM cluster, putative paralogous genes between the two clusters was identified. Genes from the minimum mitorubrin cluster covering the range between the two NR-PKS genes along with flanking genes (ranging from UA08_05975 to UA08_05994) were used in a blastx search against the putative *T. atroroseus* proteome. The five best hits from each search were outputted. Hits located within 20 kb from the putative boundaries of the MP cluster (see Chapter 5) and satisfying the criteria of e value < $1x10^{-5}$, query coverage > 70 %, amino acid similarity > 50 % were identified as putative paralogs. Similar criteria has been used by others in study of evolution of secondary metabolite clusters in fungi (Greene et al., 2014).

6.2.3 Gene tree construction and reconciliation analysis

Gene trees were constructed at protein level for the paralogous gene pairs identified above. To identify more homologs, the mitorubrin cluster gene product for each gene pair was used as query for blastp search against a local database. The database contained proteins from genome sequenced fungal species belonging to *Eurotiales*. More specifically the database contained the proteomes of *Aspergillus flavus* NRRL3357, *Aspergillus niger* CBS513.88, *Aspergillus nidulans* FGSC A4, *Aspergillus oryzae* RIB40, *Aspergillus ruber* CBS135680, *Monascus purpureus* NRRL1596, *Monascus ruber* NRRL1597, *Penicillium digitatum* PHI26, *Penicillium rubens* Wisconsin 54-1255, *Talaromyces aculeatus* ATCC10409, *Talaromyces atroroseus* IBT11181, *Talaromyces marneffei*

PM1, *T. marneffei* ATCC18224, *Talaromyces stipitatus* ATCC10500. Hits with e value < $1x10^{-5}$ were used to construct initial gene trees. The protein sequence sets were first aligned using MAFFT v6.902b with the L-INS-I option (Katoh et al., 2002; Katoh & Toh, 2008). For the PKS set of proteins only the KS domains were extracted and used for the alignment. The alignments were manually inspected and sites with more than 20% missing data were removed with trimAL v1.4 (Capella-Gutiérrez et al., 2009). Initial trees were constructed using RAxML v8.1.20 (Stamatakis, 2006) with the PROTGAMMAJTT model and 100 bootstrap replicates. The trees were manually inspected and distant sequences were eyeballed out.

The reduced set of proteins were subsequently used for phylogenetic gene tree reconstructions with the DLTRS-model as incorporated in PrIME-DLTRS created by Sjöstrand and co-workers (Sjöstrand et al., 2014). The DLTRS-model incorporates Duplication, Loss and Transfers along with Rates and provided Sequences and within a Bayesian framework constructs species tree-aware gene trees. Alignment of each set of homologous proteins was done as described above. Together with the dated species tree, the alignments were used to construct gene trees under the DLTRS-model with JTT as amino acid substitution model. The MCMC was run in three parallel chains with 2,000,000 iterations each and sampling every 500. Convergences of the MCMC runs were assessed by Tracer (Rambaut et al., 2014). The current implementation of the DLTRS-model in jPrime does not sample the posterior reconciliation; however the developers are working on getting this implemented (Sjöstrand & Arvestad, personal communication). Instead, reconciliation of the sampled gene trees to the species tree needs to be constructed using available parsimonious algorithms. Consensus trees for the individual gene trees were created by TreeAnnotator as maximum clade credibility trees combining the individual runs and with a burn in of 20%. Reconciliation analysis was hence performed using the reconciliation algorithm implemented in Notung v2.8 (Chen et al., 2000; Vernot et al., 2008; Stolzer et al., 2012). Duplication, transfer and loss rates were varied to sample a larger fraction of the parameter space.

6.3 Results and discussion

6.3.1 Mitorubrin gene cluster microsynteny and chromosomal organisation

The mitorubrin cluster was first identified in *T. marneffei* in a genome-wide PKS knockout study (Woo et al., 2012) where knockouts of Tm*pks11* and Tm*pks12* both resulted in the loss of production of mitorubrinic acid and mitorubrinol. The role of the **T**m*pks11* ortholog *mitA* in *T. atroroseus* confirmed the involvement of the cluster in the production of azaphilones belonging to the mitorubrins. Microsyntenic comparison of the mitorubrin cluster with the cluster region from other genome sequenced *Talaromyces* reveals that the cluster is present in all examined species and that the core part of the cluster possess a highly conserved synteny (Figure 6.1).

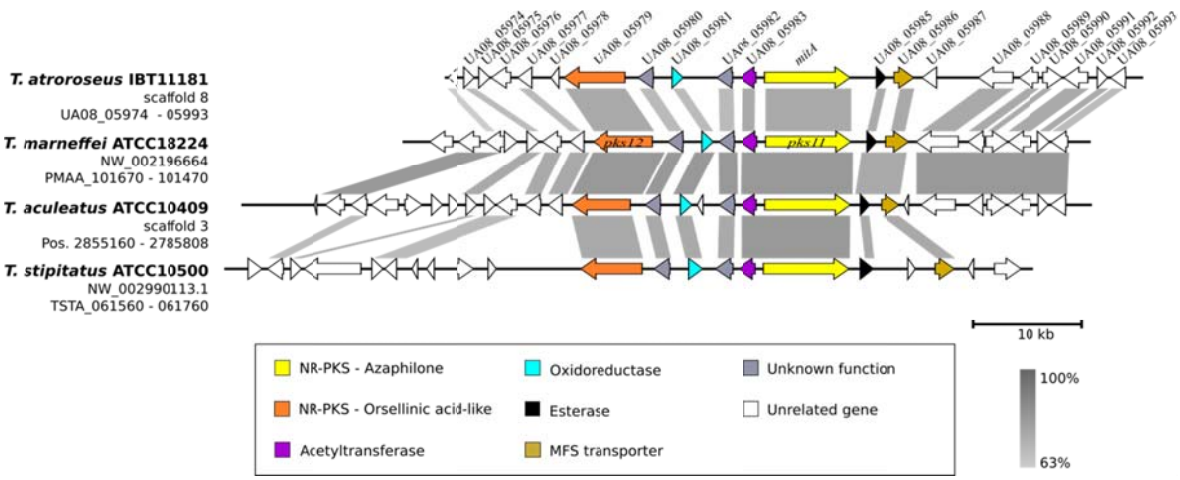


Figure 6.1. Microsynteny of the mitorubrin gene cluster across *Talaromyces* species. The mitorubrin cluster harbours two NR-PKS genes. *mitA* is responsible for the production of the azaphilone scaffold, while the UA08_05979 encoded PKS are likely to produce the orsellinic acid-like moiety in the mitorubrins. The boundaries of the mitorubrin cluster are not known hence unrelated genes are hypothetically assigned on the basis of discontinued microsynteny combined with putative functions.

Talaromyces stipitatus has so far not been reported to produce mitorubrins. The microsynteny shows that the cluster conservation is discontinued in *T. stipitatus* at both ends of the cluster. This might be some of the explanation for no observed production of mitorubrins in *T. stipitatus*. In Figure 6.1 the minimum mitorubrin gene cluster based on the discontinued synteny is assigned with putative functions, while other genes are considered putatively unrelated. However, the functional mitorubrin gene cluster might very well be larger than the genes assigned with putative functions in Figure 6.1. The extension of the cluster needs to be investigated more closely by i.e. individual knockout of individual cluster genes or by studying

differential gene expression from conditions, with and without mitorubrin production. With the previous shown cross-chemistry between the MP and mitorubrin pathway in *T. atroroseus* (Chapter 5), gene knockout studies to study the biosynthetic basis for mitorubrins should preferably be done in a *Talaromyces* strain were the MP cluster has been deleted to avoid cross-cluster effects.

6.3.2 Evolutionary scenarios of the azaphilone clusters in *Monascus* and *Talaromyces*

The MitA complementation of the lost MP PKS in *T. atroroseus* strongly indicates a tight evolutionary coherence of the PKS genes in the two azaphilone clusters. The space of possible evolutionary scenarios is in principle endless, however most scenarios are highly unlikely. Based on the prior knowledge of species relations, chemotaxonomy and gene cluster microsynteny, as discussed above, I propose three likely scenarios (Figure 6.2) for the evolution of the azaphilone clusters in *Talaromyces* and *Monascus*.

Scenario 1 – *Duplication of ancestral cluster followed by cluster modifications and losses.* In this scenario an ancestral azaphilone cluster was duplicated before the speciation event separating *Monascus* and *Talaromyces* genera. Modifications to the gene clusters, including recruitment of tailoring enzymes and other genes to the clusters, occurred before the speciation. Following the speciation the mitorubrin cluster was lost in the *Monascus* lineage. The MP cluster has undergone rearrangements in *Monascus*, *Talaromyces* or both lineages, leading to the microsynteny in the clusters seen in the extant species (See figure 5.6 in Chapter 5). Following the speciation of *T. atroroseus* from the rest of *Talaromyces* the MP PKS gene was lost in *T. atroroseus*. The driving force for the loss could be explained as the redundancy in having two PKSs producing the same or very similar azaphilone scaffold, making it more efficient to only use one of them for both pathways.

Scenario 2 – *Evolution of mitorubrin and MP* cluster *from same ancestral cluster in* Monascus *and* Talaromyces *respectively followed by horizontal gene transfer (HGT) of MP* cluster *from* Monascus *to* Talaromyces. This scenario proposes that the MP and mitorubrin gene cluster evolved from the same ancestral azaphilone cluster in a common ancestor to *Monascus* and *Talaromyces*. The MP cluster has then been horizontal transferred from *Monascus* to *Talaromyces* resulting in two azaphilone clusters in *Talaromyces*. The high degree of similarity between the MP PKS with the mitorubrin PKS could have made it more evolutionary efficient to adapt and use the

mitorubrin PKS in the biosynthesis of MPs, than to codon optimize the newly acquired *Monascus* PKS. This could in turns have resulted in the loss of the MP PKS in *T. atroroseus*.

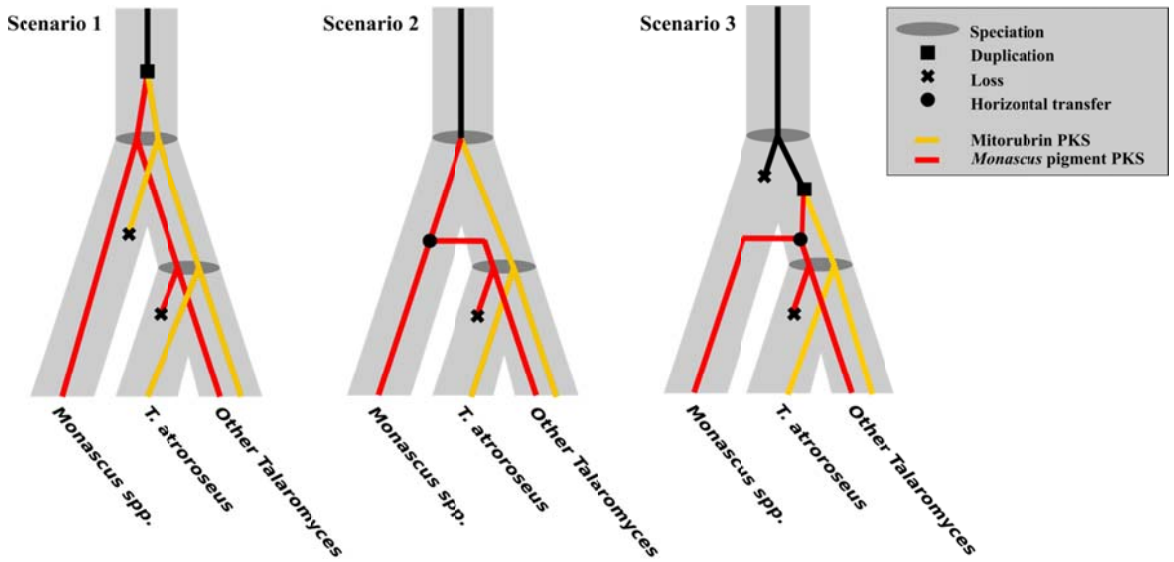


Figure 6.2. Three hypothetical scenarios of evolution of the azaphilone PKS clusters in *Talaromyces* and *Monascus* species. The scenarios is exemplified by the PKS genes were the MP PKS has been lost in *T. atroroseus*. Model 1 has a deep duplication of the azaphilone PKS followed by speciation and two losses. Model 2 involve a horizontal gene transfer (HGT) from *Monascus* to *Talaromyces* followed by the loss of the PKS gene.

Scenario 3 – *Cluster duplication within* Talaromyces *and HGT of MP cluster to* Monascus. In this scenario the two azaphilone clusters is the result of azaphilone cluster duplication in the *Talaromyces* lineage after speciation of *Monascus* and *Talaromyces*. The MP gene cluster has then been transferred horizontally from *Talaromyces* to *Monascus*. This scenario is supported by the lack of known production of either mitorubrins or MPs within other *Eurotiales,* especially the well-studied genera of *Aspergillus* and *Penicillium,* which are both in a monophyletic group together with *Monascus*. This group is separate from the monophyletic group with *Talaromyces, Rasamsonia*, *Thermomyces* and others (Houbraken et al., 2012; Yang et al., 2015).

It should be noted that the presented scenarios at this stage are purely hypothetical and not nearly covering the entire space of possible scenarios. For instance, instead of loss of the entire mitorubrin cluster in *Monascus* lineages, the cluster could simply have been partly disintegrated and scattered throughout the genome. To evaluate these different scenarios and potential other scenarios, at first homologous gene pairs, being putative paralogs between the two azaphilone clusters, were identified. The corresponding gene phylogenies were created, evaluated and reconciled to the corresponding species phylogeny.

6.3.3 Identifying putative paralogs in the two azaphilone clusters

In order to shed light on the evolutionary process behind the observed distribution of azaphilone clusters and especially to identify the most recent common ancestral azaphilone gene cluster I identified mitorubrin cluster genes with putative paralogs in the MP cluster. Blast analysis was performed using the *T. atroroseus* mitorubrin cluster genes as described in *Methods* (section 6.2.2). Apart from the two paralogous PKS genes, I found that four putative paralogous gene pairs exist between the two clusters. Until established as paralogs these will in the following be referred to as homologs.

The mitorubrin cluster gene UA08_05981 is homologous to the UA08_09246 from the MP cluster, which both are putative zinc-binding alcohol dehydrogenase containing an enoyl reductase-like domain. No function has so far been assigned to any of these gene products.

The UA08_05982 being homologous to the MP cluster UA08_09252 contains no identified functional domains. However the UA08_09252 ortholog *mpp7* from *M. purpureus* contains a NRPS condensation domain and has been characterized to be involved in the control of regioselective Knoevenagel condensation in the creation of the 2-furanone moiety of the MP scaffold (Balakrishnan et al., 2014). The lack of a 2-furanone moiety in the mitorubrins might suggest a different role of the UA08_05982 gene. UA08_05982 however also show homology to another mitorubrin cluster gene UA08_05980 which is also of unknown function.

UA08_05983 from the mitorubrin cluster is homologous to UA08_09241. Both ggene products contain a condensation domain making them putative acyltransferases. These are likely involved in the transfer of the 3-oxoacyl or orsellinic acid moieties to the azaphilone rings, respectively, in the two different azaphilone families.

The UA08_05985 is homologous to the UA08_09248 and both having putative esterase-lipase activity. They are homologous to *azaC* from the azanigerone cluster in *A. niger* (Zabala et al., 2012)*, afoE* from asperfuranone cluster in *A. nidulans* (Chiang et al., 2009) and *ctnB* from the citrinin cluster in *Monascus* species (Shimizu et al., 2007). Zabala *et al.* studied the involvement of AzaC in heterocycle formation of the bicyclic azaphilone core (Zabala et al., 2012) and excluded that hypothesis. The involvement in the azaphilone biosynthesis of these esterase-lipase enzymes remain so far unknown and should be studied further for the entire understanding of azaphilone biosynthesis.

6.3.4 Species tree and divergence dating

Time-consistency when inferring possible HGT events is an important criterion to take into account in order to rule out reconciliations accompanying HGT events that is inconsistent to the speciations going on. For time-consistent inference however a dated (relative or absolute) ultrametric species tree is required. BEAST 2.0 is a software platform employing a Bayesian framework to create divergence dating on constructed trees. For dating the tree, priors specifying evolutionary rates or calibration points are needed. As the evolutionary rates within the constructed tree is often unknown calibration points is often used in the form of fossil calibration or secondary calibration points using divergence dating inferred in other reliable studies.

In order to infer possible evolutionary scenarios for the two azaphilone pigment clusters firstly I created a dated species tree focusing on species within *Eurotiales*. It should be noted that divergence dating is tricky especially because of the lack of available fossil data for calibration within the *Eurotiales*. Since the main purpose of the divergence dating was to provide a dated species tree for use in species tree-aware gene tree reconstruction as well as in reconciliation of the gene trees to the species tree it is however more important to assign good relative dating rather than absolute dating. With this in mind it was decided to use a range of secondary calibration points from other studies in the species tree construction. Taxa outside *Eurotiales* were added to have support nodes for secondary calibration points.

The inferred species tree topology (Figure 6.3) is in agreement with the ML topology inferred by RAxML (data not shown), as well as what have been inferred by others (Hibbett et al., 2007; Samson et al., 2011; Houbraken et al., 2012; Prieto & Wedin, 2013). Within *Eurotiales* the divergence of *Talaromyces* is the first genus to diverge followed by *Penicillium* and most recent the split of *Monascus* and *Aspergillus* occurred. Within the *Talaromyces* genus clade *T. atroroseus* is the first of the sampled species to diverge with a divergence time estimation HPD95% interval of 34 to 102 Mya (Million years ago). *Talaromyces marneffei* is the last to diverge as individual species within *Talaromyces,* which could infer that being the only dimorphic opportunistic human pathogen within *Talaromyces* that these traits have been acquired rather recently. It is also worth noting that the speciation event of *Monascus purpureus* and *Monascus ruber* happened relative recent between 1.2 and 11.7 Mya indicating that these species is very closely related (Figure 6.3).

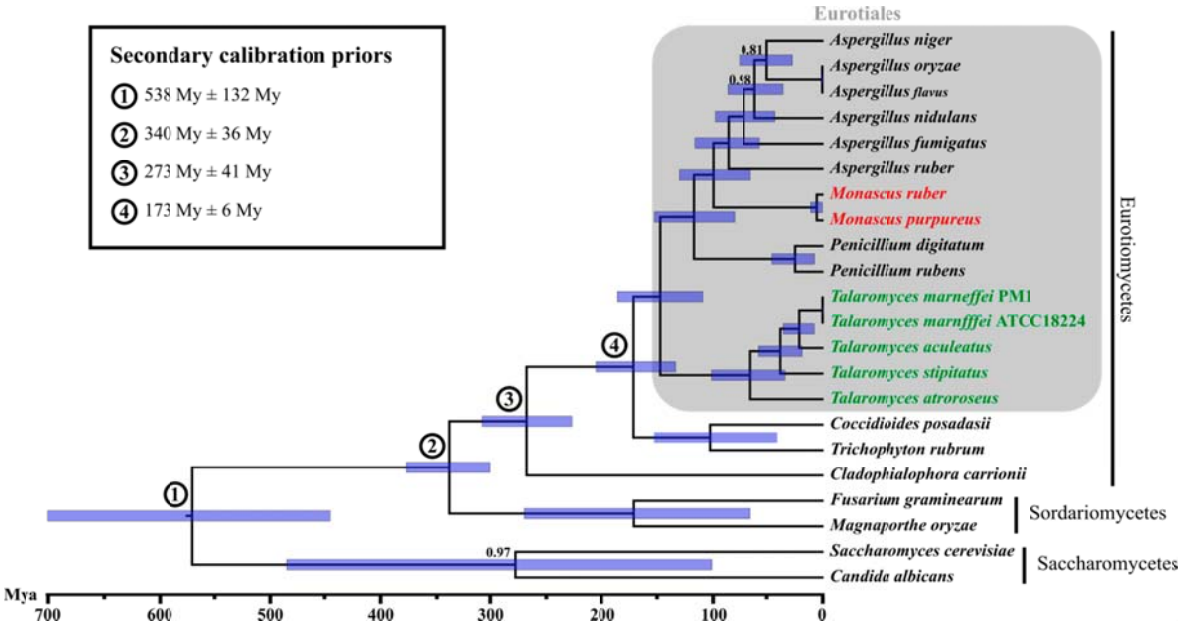


Figure 6.3. **Maximum clade credibility species tree.** Constructed with BEAST using eight unlinked genes. Each node represents the median divergence time with posterior probabilities highlighted when different from 1. The bars show the corresponding 95% HPD intervals of divergence ages. Secondary calibration points (Prieto & Wedin, 2013) is highlighted by circled numbers and applied with normal-distributed priors, see box for the corresponding 95% confidence intervals applied.

6.3.5 Species tree-aware inference of gene trees

To evaluate the evolutionary coherence of the mitorubrin and MP clusters in *Talaromyces* I constructed phylogenetic trees for genes with putative paralogs in both biosynthetic clusters as identified above. When using phylogenetics to infer evolutionary stories of genes the method for construction the gene trees can significantly change the resulting interpretation. In recent years several studies have shown that using the information from the species tree in the gene tree construction or as support to modify the tree topology post construction, can significantly improve the resulting output phylogeny (Sjöstrand et al., 2012; Wu et al., 2013; Szöllősi et al., 2014; Sjöstrand et al., 2014). The gene phylogenies was constructed with PRIME-DLTRS (Sjöstrand et al., 2014) which employs a Bayesian probabilistic framework together with models for duplication, loss and transfer events as well as a dated species tree to construct probabilistic sets of gene trees. The set of gene trees were summarized as maximum clade credibility trees and are presented in Figure 6.4-6.8 and discussed individually below. For all trees the clades with the genes from the mitorubrin and the MP cluster all reconstruct the same phylogeny, however the phylogenetic linkage between the clades differ in the different gene trees, more on this a bit later.

6.3.6 Reconciliation of gene trees

Reconciliation of the constructed gene trees were performed with Notung v2.8 (Stolzer et al., 2012), which assigns event costs of duplication, loss and transfer and reconstruct the most parsimonious reconciliation (MPR). Most reconciliation algorithms have standard DTL costs set to (D,T,L)=(2,3,1) in the case of Ranger-DTL (Bansal et al., 2012) or (1.5,3,1) in the case of Notung (Stolzer et al., 2012). However the used event costs will in some cases dramatically change the resulting reconciliation, making the assignment of event costs one of the major weaknesses of using parsimonious framework in reconciliation. To overcome this partially the DTL event costs were varied to better sample the parameter space. Since the reconciliation in this study has to do with SM cluster genes, which are known to rapidly be lost when not needed the focus by varying the event cost was to reduce loss cost and increase the transfer cost for investigation of the robustness of potentially inferred HGT events. The results from the reconciliation runs with Notung v2.8 are summarized in Table 6.1.

From the reconciliations it is clear that scenario 3 with a HGT event from *Talaromyces* to *Monascus* can be ruled out. For low transfer costs several of the trees result in reconciliations being temporally infeasible. Notung works by first finding the most parsimonious solutions and then check for feasibility. If the most parsimonious solution contains a transfer that is time- inconsistent with the species tree it will not search for temporal feasible solutions but instead report that no temporal feasible solution was found. In the following I will go through the phylogeny and reconciliation of each gene tree discussing their support for the different evolutionary scenarios.

Table 6.1. Reconciliations with Notung v2.8 (Vernot et al., 2008; Stolzer et al., 2012). For reconciled gene trees see appendix F Figure F1-F21. Evolutionary scenarios as described in Figure 6.3 have been assigned describing the reconciled evolutionary relationship between the mitorubrin and MP cluster genes. When **HGT** is added to an evolutionary scenario it means that the reconciled tree contains additional transfer events not described by the scenarios. These HGT events can be seen in the reconciled gene trees in the appendix figures.

Gene tree	Costs			# MPR	Reconciliation score	Evolutionary scenario	Reconciled trees
	Duplication	Transfer	Loss				
MitA	2	3	1	-*			
	2	4	1	-*			
	2	5	1	-*			
	2	7	1	1	40	**1**	Figure F.1
	2	9	1	1	40	**1**	Figure F.1
	2	11	1	1	40	**1**	Figure F.1
	4	6	1	-*			
	4	8	1	-*			
	4	10	1	1	54	**1**	Figure F.1
	4	14	1	1	54	**1**	Figure F.1
	6	10	1	-*			
	6	14	1	1	68	**1**	Figure F.1
UA08_05983 / UA08_09241	2	3	1	4	24	**2/3 + HGT**	Figure F.2-F.5
	2	4	1	3	29	**2+HGT**	Figure F.6-F.8
	2	5	1	1	31	**2 +HGT**	Figure F.8
	2	7	1	2	35	**1, 2 + HGT**	Figure F.8-F.9
	2	9	1	1	43	**1**	Figure F.9
	2	11	1	1	40	**1**	Figure F.9
	4	6	1	1	41	**2 + HGT**	Figure F.8
	4	8	1	1	45	**2 + HGT**	Figure F.8
	4	10	1	1	47	**1**	Figure F.9
	4	14	1	1	47	**1**	Figure F.9
	6	10	1	1	57	**2 + HGT**	Figure F.8
	6	14	1	1	59	**1**	Figure F.9
UA08_05981 / UA08_09246	2	3	1	-*			
	2	4	1	-*			
	2	5	1	-*			
	2	7	1	1	53	**1**	Figure F.10
	2	9	1	1	53	**1**	Figure F.10
	2	11	1	1	53	**1**	Figure F.10
	4	6	1	-*			
	4	8	1	6	73	**1, 2 + HGT**	Figure F.10-F.15
	4	10	1	1	73	**1**	Figure F.10
	4	14	1	1	73	**1**	Figure F.10
	6	10	1	6	93	**1, 2 + HGT**	Figure F.10-F.15
	6	14	1	1	93	**1**	Figure F.10
UA08_05985 / UA08_09248	2	3	1	2	40	**2 + HGT**	Figure F.16-F.17
	2	4	1	-*			
	2	5	1	-*			
	2	7	1	2	58	**1, 1 + HGT**	Figure F.18-F.19
	2	9	1	1	58	**1**	Figure F.19
	2	11	1	1	58	**1**	Figure F.19
	4	6	1	-*			
	4	8	1	-*			
	4	10	1	1	82	**1**	Figure F.19
	4	14	1	1	82	**1**	Figure F.19
	6	10	1	-*			
	6	14	1	1	106	**1**	Figure F.19
UA08_05980 / UA08_09252 / UA08_05982 /	2	3	1	1	17	**2**	Figure F.20
	2	4	1	1	18	**2**	Figure F.20
	2	5	1	1	19	**2**	Figure F.20
	2	7	1	1	21	**2**	Figure F.20
	2	9	1	2	23	**1, 2**	Figure F.20-F.21
	2	11	1	1	23	**1**	Figure F.21
	4	6	1	1	24	**2**	Figure F.20
	4	8	1	1	26	**2**	Figure F.20
	4	10	1	1	28	**2**	Figure F.20
	4	14	1	1	31	**1**	Figure F.21
	6	10	1	1	32	**2**	Figure F.20
	6	14	1	1	36	**2**	Figure F20

6.3.7 Phylogenetic interpretation of gene tree

MitA gene tree - The inferred phylogeny of the azaphilone NR-PKS homologues to MitA is shown in Figure 6.4. The mitorubrin and MP PKSs groups in two clades only separated by the *Aspergillus niger* PKS AzaA, which is the azaphilone PKS involved in the biosynthesis of azanigerones. Azanigerones have structural similarity to the mitorubrins but instead of having an orsellinic acid moiety the attached side group is a 2,4-dimethylhexanoyl chain (Zabala et al., 2012). Comparing the presented phylogeny with the species tree strongly supports Scenario 1 involving a duplication of the PKS gene existing in a common ancestor of *Monascus* and *Talaromyces* and the MP PKS has following been lost in the *Aspergillus* and *Penicillium* lineages. The mitorubrin PKS has been lost in *Monascus, Penicillium* and most Aspergilli. However in *A. niger* it has evolved to AzaA, which produces the azaphilone core the azanigerone compounds. This interpretation of the gene phylogenies is also supported by the Notung reconciliation where all DTL combinations giving temporal feasible solution have Scenario 1 as the MPR.

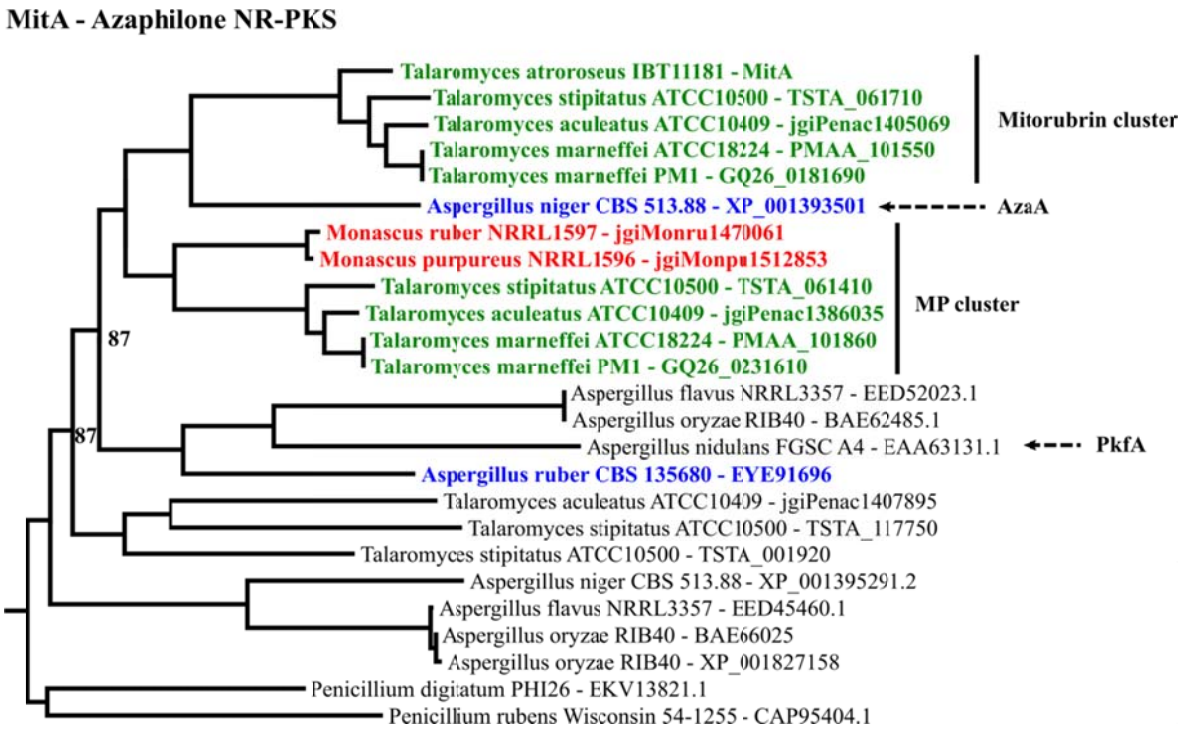


Figure 6.4. Maximum clade credibility tree of mitorubrin / MP PKS homologs created with PrIME-DLTRS. Posterior probabilities are highlighted on nodes when below 0.95. The clades with PKSs from the known mitorubrin and MP gene clusters are highlighted and important tip names are coloured according to their genus: *Talaromyces* (Green), *Monascus* (Red), *Aspergillus* (Blue).

UA08_05983 / UA08_09241 gene tree - Scenario 1 is also strongly supported by the phylogeny of the acyltransferase homologs UA08_05983 / UA08_09241 (Figure 6.5). The phylogenetic grouping of the mitorubrin and MP cluster proteins is the same as seen for the MitA phylogeny, except that *T. atroroseus* harbours both the mitorubrin and MP cluster acyltransferase in contrast to the MitA where the MP PKS was lost. Furthermore two *A. niger* genes (highlighted in blue) groups in a monophyletic group with the mitorubrin genes one of them being AzaD from the *aza* cluster meaning that the *azaD* gene has duplicated after the speciation of *Aspergillus* and *Talaromyces*.

Reconciliations with low transfer costs favour Scenario 2 where the *mon* acyltransferase has been horizontal transferred to *Talaromyces*. This reconciliation model, which can be found in Figure F8 also infer a transfer of the mitorubrin cluster acyltransferase from a *Talaromyces* ancestor to *A. niger*. When the transfer costs are raised the most parsimonious reconciliation only involves duplications and losses and is consistent with Scenario 1.

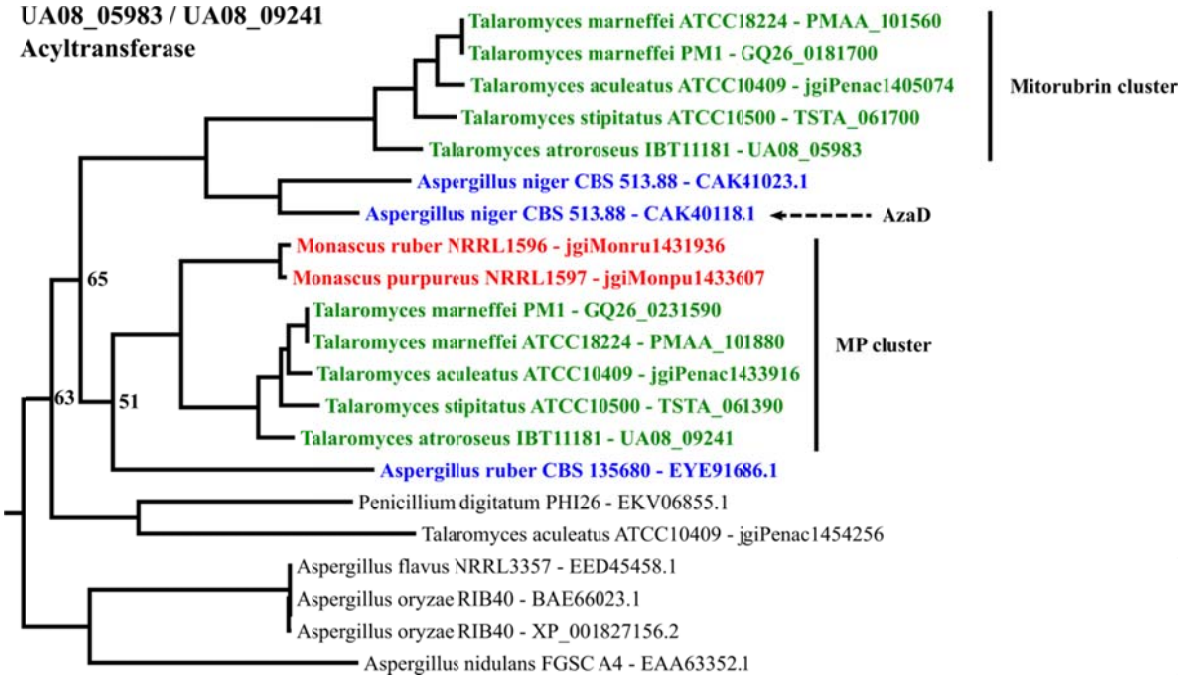


Figure 6.5. Maximum clade credibility tree of UA08_05983 / UA08_09241 homologs created with PrIME-DLTRS. Posterior probabilities are highlighted on nodes when below 0.95. The clades with PKSs from the known mitorubrin and MP gene clusters are highlighted and important tip names are coloured according to their genus: *Talaromyces* (Green), *Monascus* (Red), *Aspergillus* (Blue).

UA08_05981/UA08_09246 gene trees - The reconciliation of both the phylogenies concerning the gene pairs UA08_05981/UA08_09246 and UA08_05985/UA08_09248 support scenario 1 of duplication and loss when temporal feasible reconciliations were present. In the UA08_05981/UA08_09246 tree of putative enoyl reductases (Figure 6.6) the three *Aspergillus* gene products, including AzaJ, nest between the clade containing the mitorubrin cluster enoyl reductases and another big clade with enoyl reductases from *Talaromyces*. This would normally indicate the presence of a HGT event from *Talaromyces* to *Aspergillus* but this is however not supported by the reconciliation and by the other gene trees. Instead duplication of the mitorubrin cluster enoyl reductase has occurred before the split of *Talaromyces* and *Aspergillus* and following the duplication extensive loss of enoyl reductases has probably occurred in *Aspergillus* and *Penicillium* species.

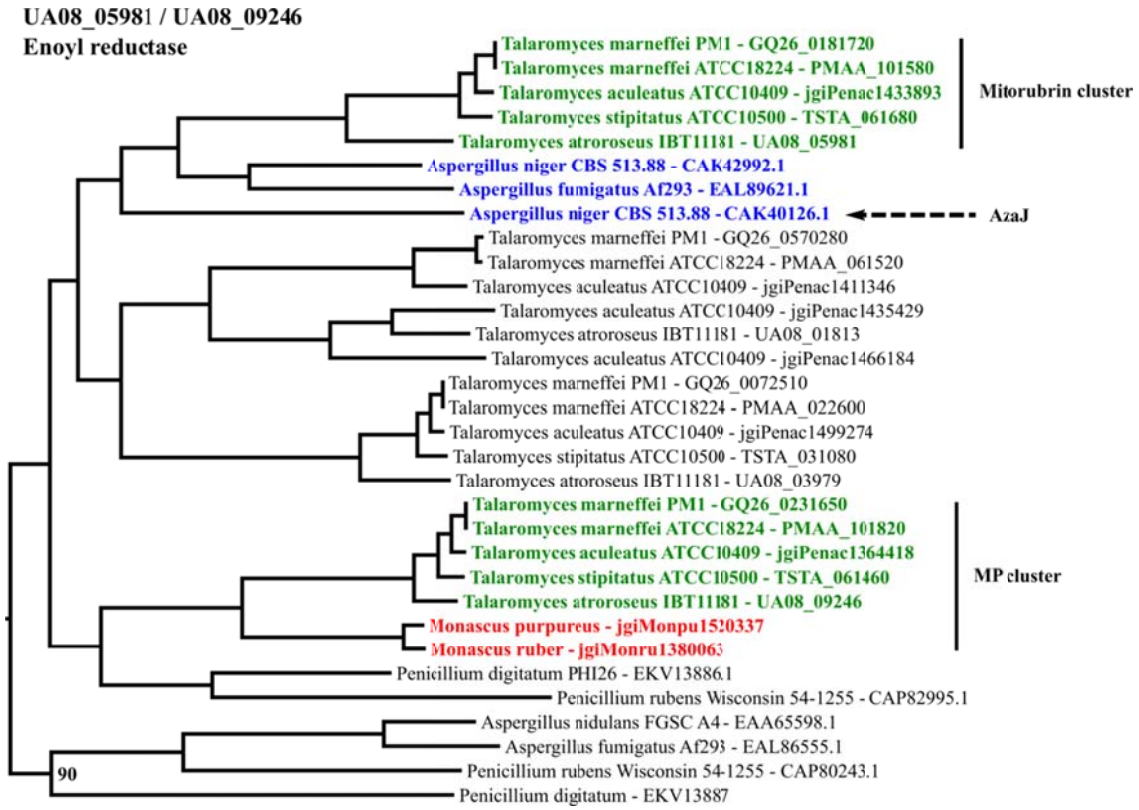


Figure 6.6. Maximum clade credibility tree of UA08_05981 / UA08_09249 homologs created with PrIME-DLTRS. Posterior probabilities are highlighted on nodes when below 0.95. The clades with PKSs from the known mitorubrin and MP gene clusters are highlighted and important tip names are coloured according to their genus: *Talaromyces* (Green), *Monascus* (Red), *Aspergillus* (Blue).

UA08_05985/UA08_09248 gene tree - The interpretation of the UA08_05985/ UA08_09248 tree is a bit more complex since this is the only tree where an *aza* cluster gene product is present but does not group together with the mitorubrin cluster gene products. However as several of the nodes have very low probabilities this might not reflect the true phylogeny. Instead a number of other *Aspergillus* putative esterases group closely with the mitorubrin cluster gene products among these AfoC, which comes from the asperfuranone cluster in *A. nidulans*. Furthermore it is worth to note that a *Monascus purpureus* gene also group with together with the *mit* genes. This strongly supports that the mitorubrin and MP cluster originate from a deep duplication of an ancestral cluster, as in scenario 1. The mitorubrin cluster has been heavily reduced in *Monascus* but the putative esterase gene is retained in *M. purpureus*.

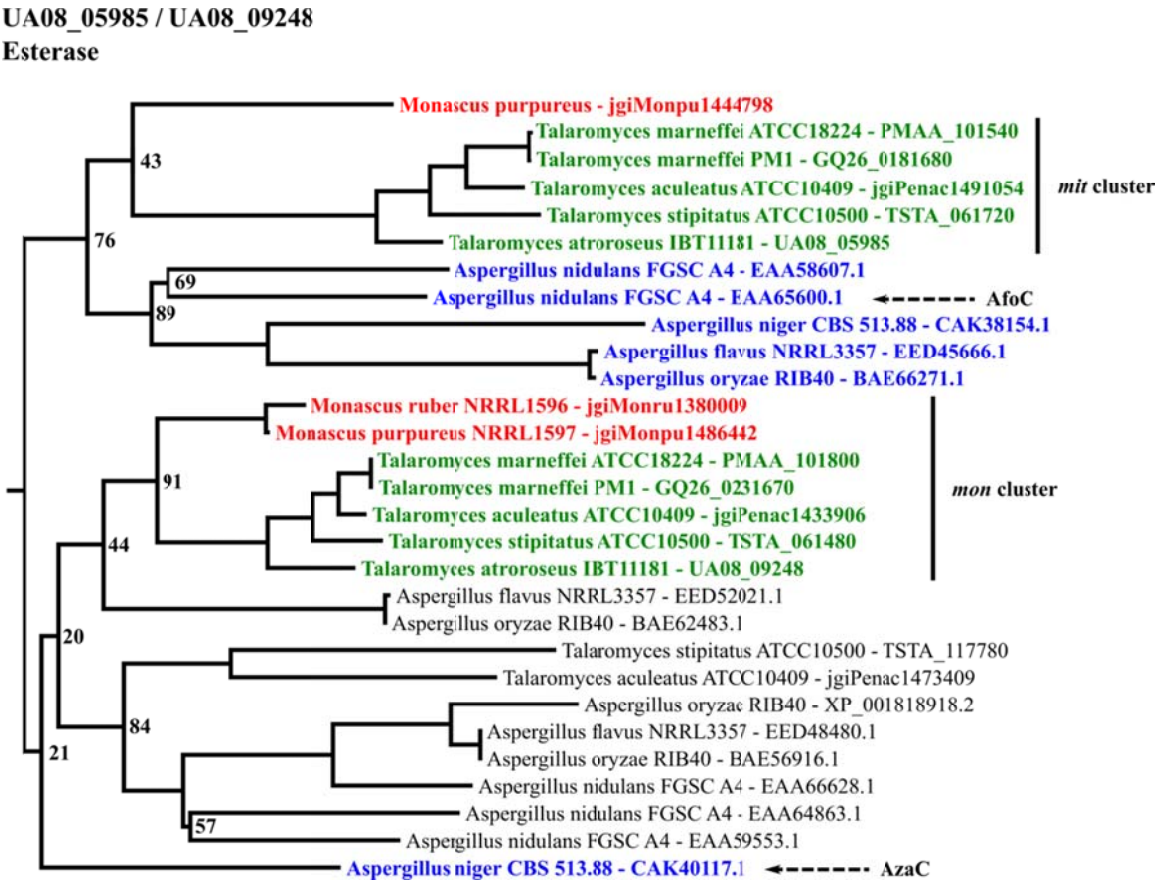


Figure 6.7. Maximum clade credibility tree of UA08_05983 / UA08_09241 homologs created with PrIME-DLTRS. Posterior probabilities are highlighted on nodes when below 0.95. The clades with PKSs from the known mitorubrin and MP gene clusters are highlighted and important tip names are coloured according to their genus: *Talaromyces* (Green), *Monascus* (Red), *Aspergillus* (Blue).

UA08_05982/UA08_09252/UA08_05980 gene tree - The last set of homologous genes contained two genes identified in the mitorubrin cluster i.e. UA08_05982 and UA08_05980. From the identification of homologous genes between the mitorubrin and MP cluster previously, UA08_05982 had a homolog in the MP cluster i.e. encoded by *mpp7*. The phylogeny shown in Figure 6.8 however reveals that UA08_05980 also have a homolog in the MP cluster in *Monascus* (the upper clade: jgiMonru1175953 and jgiMonpu1494348). This MP homolog, encoded by *mpp8* referring to the gene naming introduced by Balakrishnan *et al.* (Balakrishnan et al., 2013; Balakrishnan et al., 2014), is located in the MP cluster in *Monascus,* two genes downstream of *mpp7*. The phylogenetic relationship of *mpp8* and the subclade containing UA08_05980 is equal to the relationship between the MP cluster ortholog *mpp7* and UA08_05982 (Bottom clade, Figure 6.8).

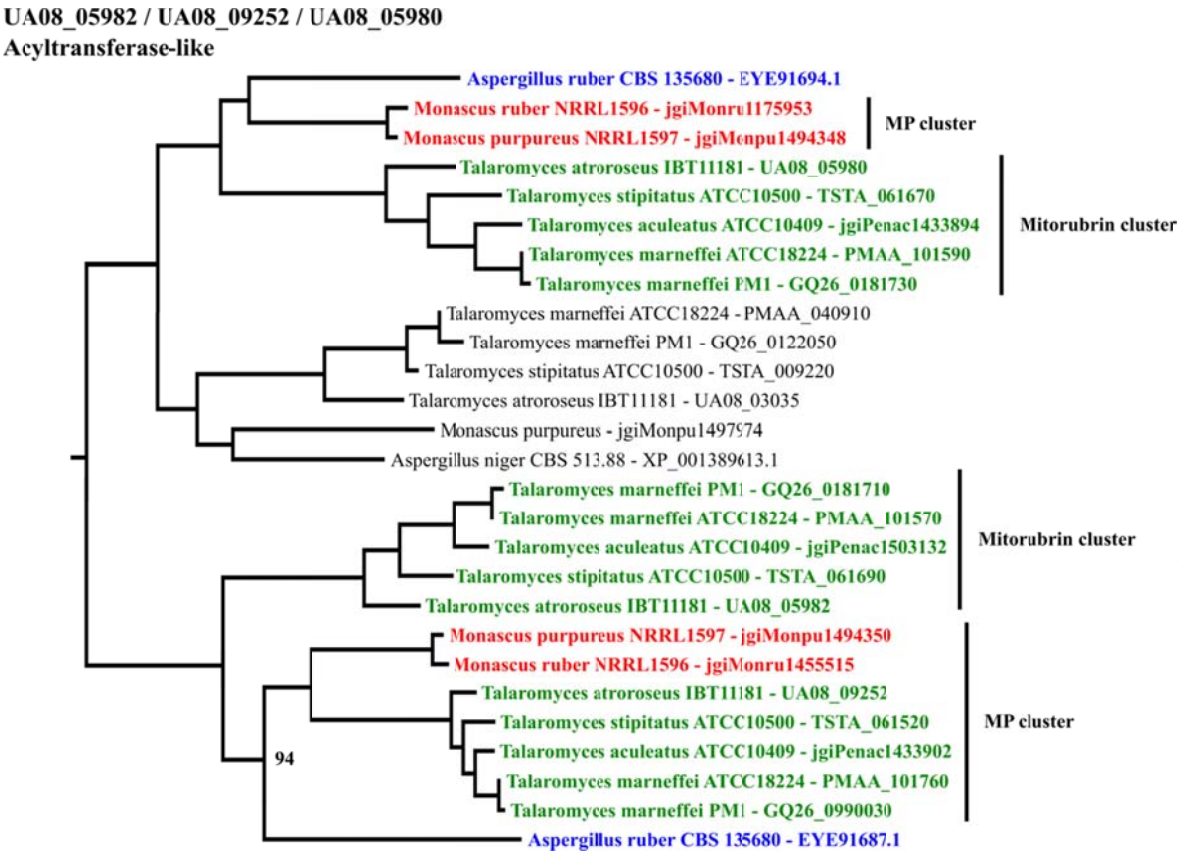


Figure 6.8. Maximum clade credibility tree of UA08_05982 / UA08_09252 / UA08_05980 homologs created with PrIME-DLTRS. Posterior probabilities are highlighted on nodes when below 0.95. The clades with PKSs from the known mitorubrin and MP gene clusters are highlighted and important tip names are coloured according to their genus: *Talaromyces* (Green), *Monascus* (Red), *Aspergillus* (Blue).

The reconciliation with low transfer cost support Scenario 2, however the inferred transfer event is only occurring in the lower clade. The two *A. ruber* proteins with accession EYE91687.1 and EYE91694.1 groups respectively in the lower and the upper clade in the phylogenetic tree (Figure 6.8). However in the upper clade the *A. ruber* protein form a monophyletic clade with the *Monascus* proteins, while it in the lower clade branch outside the monophyletic clade with the MP cluster genes. This difference in positioning is the reason for the inference of HGT event only in the lower clade. While this tree cannot rule out the occurrence of a HGT event, it is more likely that the entire phylogeny is explained by duplication and losses. Especially the similar phylogenetic relationship in the upper and the bottom clade supports this. The *mpp8* / UA08_05980 clade is hence a result of a gene duplication occurring before the putative duplication of the azaphilone cluster. After the duplication of the azaphilone cluster as exemplified in the lower clade by *mpp7* and UA08_05982 the fate of *mpp8* and UA08_05980 is a bit different. While *Monascus* species has not retained the resulting mitorubrin cluster the *mpp8* was retained and recruited or kept in the MP cluster. In *Talaromyces* species both azaphilone clusters was retained following the cluster duplication and UA08_05980 was recruited or kept in the mitorubrin cluster.

The genes for the two discussed *A. ruber* proteins are located next to each other and 9 and 10 genes respectively downstream of the PKS gene encoding EYE91696. This PKS however cluster outside the clade with mitorubrin and MP PKSs (Figure 6.4) and instead show closer relationship to the *A. nidulans* PkfA, which has been linked to the production of the small aromatic compound orsellinaldehyde (Ahuja et al., 2012). It was later shown that the orsellinaldehyde is a precursor in the biosynthetic pathway to Aspernidine A (Yaegashi et al., 2013).

In the phylogeny of the acyltransferase (Figure 6.5) *A. ruber* has retained the copy of the acyltransferase of the MP cluster (EYE91686.1), while it has been lost in all other examined Aspergilli. The low clade probability of 51 percent at the node linking the *A. ruber* to the MP cluster clade suggests that alternative explanations to the origin of the *A. ruber* acyltransferase might exist. However the phylogeny observed is similar to what is observed for in the lower clade of the UA08_05982/UA08_09252/UA08_05980 phylogenetic tree.

The phylogenetic differences of the *A. ruber* genes in the three different trees inflicts that the true evolutionary relationship between the *A. ruber* gene cluster containing the EYE91696 gene and the azaphilone clusters in *Talaromyces* and *Monascus* remains elusive.

6.3.8 Further support for duplication-loss evolutionary scenario

A closer investigation of the chromosomal organisation of the azaphilone clusters in *Talaromyces* reveals that the mitorubrin cluster is located less than 50 kb away from the MP cluster in all *Talaromyces* species except *T. atroroseus* (Figure 6.9). In *T. atroroseus* the two azaphilone clusters are located towards the end of two separate scaffolds, scaffold_8 and scaffold_19. It is likely that these scaffolds are linked on the chromosome in an end to end manner giving the same relative chromosomal organization of the azaphilone clusters as is seen in the other *Talaromyces*. I have tried to link the two scaffolds together by PCR but I have however so far not been successful. The least labour-intensive option for linking scaffolds in the *T. atroroseus* genome together to show if scaffold_8 and scaffold_19 is truly linked would be to do additional sequencing using rapid evolving sequencing technology with long read lengths e.g. PacBio sequencing combined with improved algorithms and pipelines for genome assembly as described by English *et al.* (English et al., 2012).

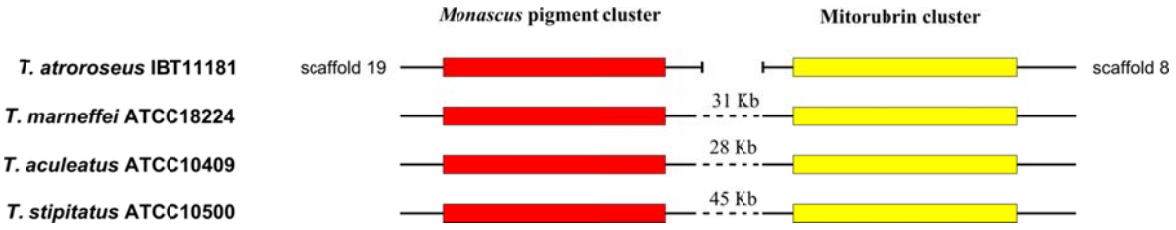


Figure 6.9. Chromosomal organisation of azaphilone gene clusters in *Talaromyces*. The two azaphilone clusters for the biosynthesis of MPs and mitorubrins are located in close proximity to each other. In *T. atroroseus* the two clusters are however located at the very end of two separate genomic scaffolds.

The close chromosomal arrangement of the two azaphilone clusters in *Talaromyces* is a strong support for the duplication-loss history in scenario 1, since such tandem rearrangement are known to be one of the possible outcomes of gene duplications by segmental repeats (Schacherer et al., 2005). Furthermore, the NR-PKS and acyltransferase in both clusters are located tail to tail, but with an inserted oxidoreductase in the MP cluster (See Figure 5.6 and 6.1). This tail to tail arrangement is additional support for the duplication loss scenario.

To sum up the evaluated gene phylogenies as well as reconciliations points towards scenario 1. Under the likely assumption that the homologous gene pairs evaluated share the same evolutionary fate I thereby conclude that an old duplication event of multiple genes from an ancestral azaphilone cluster resulted in two paralogous azaphilone clusters in the common *Eurotiales* ancestor. These two clusters evolved along different paths resulting in the MP and the mitorubrin cluster present in extant species. The mitorubrin cluster has been lost

in *Monascus* lineages while in *A. niger* it evolved into the azanigerone cluster. The MP cluster has been retained in both *Monascus* and *Talaromyces* species but has apparently been lost in examined *Aspergillus* and *Penicillium* species. In *T. atroroseus* the key enzyme in the MP cluster, i.e. the MP PKS has been lost as described in Chapter 5.

6.3.9 Loss of azaphilone cluster in *Aspergillus* and *Penicillium*

The proposed evolutionary origin of the azaphilone clusters means that the cluster duplication has happened before the speciation of *Monascus* and *Talaromyces* and probably long time before the root of the *Eurotiales*. This finding inflicts that the species rich genera of *Aspergillus* and *Penicillium* contained ancestral *mon*- and *mit*-like clusters. The azaphilone clusters has been extensively lost in many *Aspergillus* lineages, however the mitorubirn-like cluster has been retained in *A. niger* and evolved into the azanigerone cluster. This strongly supports the proposed duplication/loss model.

An unknown *Aspergillus* species has previously been shown to produce the azaphilone pigment Sch 1385568 (Yang et al., 2009), which show remarkable structural similarities with the mitorubrins but differs in having the orsellinic acid attached to the C8-linked oxygen (Figure 6.10). A similar linkage is found in (+)-mitorubrinic acid B produced by *Talaromyces funiculosus* (Formerly *Penicillium funiculosum)* (Natsume et al., 1985). This strongly suggest that the mitorubrin cluster is retained in the unknown *Aspergillus sp.* but the putative acyltransferase attaching the orsellinic acid moiety to the azaphilone scaffolds might have evolved to change specificity towards the C8-linked oxygen for attachment site.

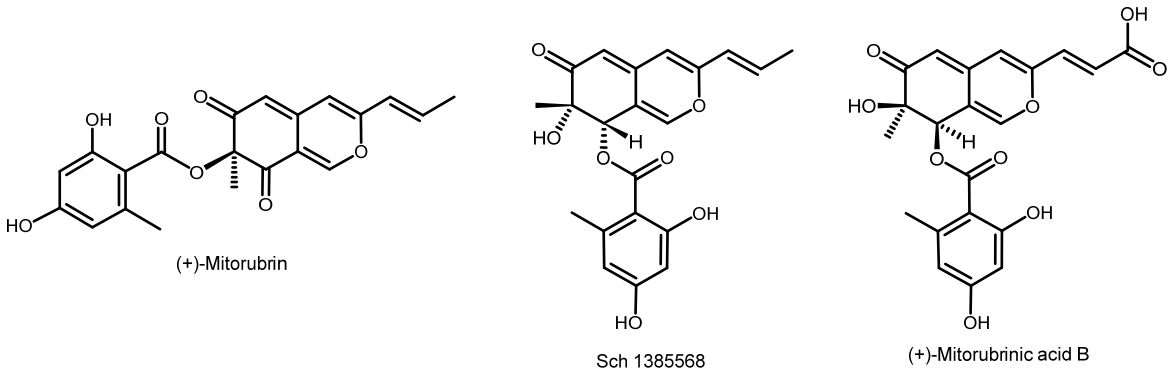


Figure 6.10. Structures of (+)-Mitorubrin, Sch 1385568 and (+)-Mitorubrinic acid. The structural similarity suggests that these two compounds are produced by similar gene clusters evolved from the same ancestral mitorubrin gene cluster from the root of *Eurotiales*.

Production of the mycotoxin citrinin, which also belongs to the azaphilones is also linked to species belonging to *Aspergillus* and *Penicillium*. Apart from the azaphilones linked to genes

within *Eurotiales* several species in the *Penicillium* genus are known to produce azaphilones, especially the production of sclerotiorin, isochromophilones and more by species such as *P. sclerotiorum* and *P. multicolor* are known (Osmanova et al., 2010). This suggests that at least some *Penicillium* species has retained an azaphilone gene cluster.

Altogether the above examples of azaphilones in *Aspergillus* and *Penicillium* suggest that the loss or retention of azaphilone clusters is far more complex than what has been addressed in the current study. Ongoing sequencing efforts of fungal *Aspergillus* genomes such as the "*Aspergillus* whole-genus sequencing project" will in the near future provide more information on whether remnants of the mitorubrin gene cluster has been retained in other *Aspergillus* species.

6.3.10 The loss of *Monascus* pigment PKS in *T. atroroseus*

The loss of the MP PKS in *T. atroroseus* IBT11181 is truly intriguing since *T. atroroseus* IBT11181 is one of the best known producers of MPs. Gene loss is part of the evolutionary mechanism leaving imprints on the present distribution of genes in fungal species. One of the driving forces for such gene loss is redundancy in gene function.

The PKS loss in *T. atroroseus* seems to have happened from a functional gene cluster according to the above discussed evolutionary model and have hence not happened recently after gene duplication. Thus, the driving force for the gene loss might be a bit different. It is clear that redundancy in gene function between the MP PKS and the mitorubrin PKS has existed in *T. atroroseus* since the mitorubrin PKS complements the loss of the MP PKS. However, what is the advantage in keeping the mitorubrin PKS over the MP PKS when the fungi under laboratory conditions is producing significantly larger amounts of the MP biosynthetic compounds? I believe that the loss of the MP PKS have created a competitive advantage for the fungus in its natural habitat. The most likely explanation is that the mitorubrin PKS evolved to be a more favourable PKS for MP production. This could be in terms of more efficient regulation resulting in higher or more rapid expression of the mitorubrin PKS, mutations hampering the function of the MP PKS making the mitorubrin PKS an advantageous alternative or slightly different array of compounds created by the mitorubrin PKS potentially having a competitive advantage for the fungi.

The group of mitorubrins include compounds such as mitorubrinic acid, mitorubrinal and mitorubrin acetate, which all have different functional groups at the polyketide backbone end. These variations have, to my knowledge, not been observed in MP produced by *Monascus*

species. However the *T. atroroseus* produced compound PP-V as well as the novel compound N-glutamyl monascorubraminic acid (from Chapter 4) both has a carbocyclic acid group at the polyketide end as seen in mitorubrinic acid. I speculate that this might be a feature that is uniquely linked to the azaphilones from the mitorubrin cluster. This can potentially be explained by starter unit promiscuity of the SAT domain in the mitorubrin PKS or by mitorubrin cluster tailoring enzymes converting the methyl-end of the polyketide to a carboxylic acid.

In *T. atroroseus* the usage of azaphilone precursors from the mitorubrin cluster in the biosynthesis of MPs could then have led to the production PP-V and N-glutamyl monascorubraminic acid and potentially more of the unknown red MPs discussed in Chapter 4. This enhanced product range might have given the fungi an advantage by the loss of the original MP PKS.

6.3.11 Species tree

Phylogenetic inference by species tree-aware methods as well as the reconciliation is highly affected by the topology of the used species tree, and it is paramount to use resolved species tree with high support. The species tree constructed in this study, however, is well supported by the phylogeny inferred in several other studies and is hence believed to be the true phylogeny of the species used in the analysis.

6.4 Conclusion

In conclusion, the present study show strong evidence that the mitorubrin and MP cluster in *Eurotiales* is a result of a deep cluster duplication involving at least the azaphilone NR-PKS, the acyltransferase and the acyltransferase-like gene UA08_05982/UA08_09252 and potentially also an enoyl reductase and an esterase. More genes might have been present in the cluster at the duplication event and since then been lost from one or both clusters as part of the adaptive evolution. Since the duplication event the two gene clusters has evolved by a combination of loss of functions, recruiting new enzyme functions and potential neofunctionalization of existing genes to the clusters present in extant species. Among the investigated Aspergilli the *aza* cluster in *A. niger* evolved divergently from the mitorubrin cluster after the duplication of the ancestral azaphilone cluster.

6.5 References

Ahuja, M., Chiang, Y.-M., Chang, S.-L., Praseuth, M.B., Entwistle, R., et al., **2012**. Illuminating the diversity of aromatic polyketide synthases in *Aspergillus nidulans*. *Journal of the American Chemical Society*, **134**: 8212–8221. doi:10.1021/ja3016395

Balakrishnan, B., Chen, C.-C., Pan, T.-M. & Kwon, H.-J., **2014**. Mpp7 controls regioselective Knoevenagel condensation during the biosynthesis of *Monascus* azaphilone pigments. *Tetrahedron Letters*, **55**: 1640–1643. doi:10.1016/j.tetlet.2014.01.090

Balakrishnan, B., Karki, S., Chiu, S.-H., Kim, H.-J., Suh, J.-W., et al., **2013**. Genetic localization and in vivo characterization of a *Monascus* azaphilone pigment biosynthetic gene cluster. *Applied Microbiology and Biotechnology*. doi:10.1007/s00253-013-4745-9

Bansal, M.S., Alm, E.J. & Kellis, M., **2012**. Efficient algorithms for the reconciliation problem with gene duplication, horizontal transfer and loss. *Bioinformatics (Oxford, England)*, **28**: i283–91. doi:10.1093/bioinformatics/bts225

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., et al., **2014**. BEAST 2: A software platform for bayesian evolutionary analysis. *PLoS Computational Biology*, **10**: 1–6. doi:10.1371/journal.pcbi.1003537

Brown, D.W., Butchko, R. a E., Baker, S.E. & Proctor, R.H., **2012**. Phylogenomic and functional domain analysis of polyketide synthases in *Fusarium*. *Fungal Biology*, **116**: 318–331. doi:10.1016/j.funbio.2011.12.005

Buechi, G., White, J.D. & Wogan, G.N., **1965**. The structures of mitorubrin and mitorubrinol. *Journal of the American Chemical Society*, **87**: 3484–3489.

Capella-Gutiérrez, S., Kauff, F. & Gabaldón, T., **2014**. A phylogenomics approach for selecting robust sets of phylogenetic markers. *Nucleic Acids Research*, **42**: e54. doi:10.1093/nar/gku071

Capella-Gutiérrez, S., Silla-Martínez, J.M. & Gabaldón, T., **2009**. TrimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, **25**: 1972–1973. doi:10.1093/bioinformatics/btp348

Chen, K., Durand, D. & Farach-Colton, M., **2000**. NOTUNG: a program for dating gene duplications and optimizing gene family trees. *Journal of Computational Biology*, **7**: 429–447. doi:10.1089/106652700750050871

Chiang, Y.M., Szewczyk, E., Davidson, A.D., Keller, N., Oakley, B.R., et al., **2009**. A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *Journal of the American Chemical Society*, **131**: 2965–2970. doi:10.1021/ja8088185

English, A.C., Richards, S., Han, Y., Wang, M., Vee, V., et al., **2012**. Mind the gap: Upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS ONE*, **7**: 1–12. doi:10.1371/journal.pone.0047768

Frisvad, J.C., Filtenborg, O., Samson, R.A. & Stolk, A.C., **1990**. Chemotaxonomy of the genus *Talaromyces*. *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology*, **57**: 179–189. doi:10.1007/BF00403953

Greene, G.H., McGary, K.L., Rokas, A. & Slot, J.C., **2014**. Ecology drives the distribution of specialized tyrosine metabolism modules in fungi. *Genome Biology and Evolution*, **6**: 121–132. doi:10.1093/gbe/evt208

Heled, J. & Drummond, A.J., **2012**. Calibrated tree priors for relaxed phylogenetics and divergence time estimation. *Systematic Biology*, **61**: 138–149. doi:10.1093/sysbio/syr087

Hibbett, D.S., Binder, M., Bischoff, J.F., Blackwell, M., Cannon, P.F., et al., **2007**. A higher-level phylogenetic classification of the Fungi. *Mycological Research*, **111**: 509–547. doi:10.1016/j.mycres.2007.03.004

Houbraken, J., Spierenburg, H. & Frisvad, J.C., **2012**. *Rasamsonia*, a new genus comprising thermotolerant and thermophilic *Talaromyces* and *Geosmithia* species. *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology*, **101**: 403–421. doi:10.1007/s10482-011-9647-1

Katoh, K., Misawa, K., Kuma, K. & Miyata, T., **2002**. MAFFT : a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acid Research*, **30**: 3059–3066.

Katoh, K. & Toh, H., **2008**. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, **9**: 286–98. doi:10.1093/bib/bbn013

Keller, N.P. & Hohn, T.M., **1997**. Metabolic pathway gene clusters in filamentous fungi. *Fungal Genetics and Biology*, **21**: 17–29. doi:10.1006/fgbi.1997.0970

Lin, S., Yoshimoto, M., Lyu, P.-C., Tang, C.-Y. & Arita, M., **2012**. Phylogenomic and domain analysis of iterative polyketide ´synthases in *Aspergillus* species. *Evolutionary Bioinformatics*, **8**: 373–387. doi:10.4137/EBO.S9796

Natsume, M., Takahashi, Y. & Marumo, S., **1985**. (-)-Mitorubrinic acid, a morphogenic substance inducing chlamydospore-like cells, and its related new metabolite, (+)-mitorubrinic acid-B, isolated from *Penicillium funiculosum*. *Agricultural and Biological Chemistry*, **49**: 2517–2519.

Osmanova, N., Schultze, W. & Ayoub, N., **2010**. Azaphilones: a class of fungal metabolites with diverse biological activities. *Phytochemistry Reviews*, **9**: 315–342. doi:10.1007/s11101-010-9171-3

Prieto, M. & Wedin, M., **2013**. Dating the Diversification of the Major Lineages of Ascomycota (Fungi). *PLoS ONE*, **8**: e65576. doi:10.1371/journal.pone.0065576

Rambaut, A., Suchard, M., Xie, D. & Drummond, A.J., **2014**. Tracer v1.6. Available from http://beast.bio.ed.ac.uk/Tracer

Samson, R.A., Yilmaz, N., Houbraken, J., Spierenburg, H., Seifert, K.A., et al., **2011**. Phylogeny and nomenclature of the genus *Talaromyces* and taxa accommodated in *Penicillium* subgenus *Biverticillium*. *Studies in Mycology*, **70**: 159–183. doi:10.3114/sim.2011.70.04

Schacherer, J., de Montigny, J., Welcker, A., Souciet, J.L. & Potier, S., **2005**. Duplication processes in *Saccharomyces cerevisiae* haploid strains. *Nucleic Acids Research*, **33**: 6319–6326. doi:10.1093/nar/gki941

Shimizu, T., Kinoshita, H., Ishihara, S., Sakai, K., Nagai, S., et al., **2005**. Polyketide synthase gene responsible for citrinin biosynthesis in *Monascus purpureus*. *Applied and Environmental Microbiology*, **71**: 3453–3457.

Shimizu, T., Kinoshita, H. & Nihira, T., **2007**. Identification and in vivo functional analysis by gene disruption of ctnA, an activator gene involved in citrinin biosynthesis in Monascus purpureus. *Applied and Environmental Microbiology*, **73**: 5097–5103.

Sjöstrand, J., Sennblad, B., Arvestad, L. & Lagergren, J., **2012**. DLRS: Gene tree evolution in light of a species tree. *Bioinformatics*, **28**: 2994–2995. doi:10.1093/bioinformatics/bts548

Sjöstrand, J., Tofigh, A., Daubin, V., Arvestad, L., Sennblad, B., et al., **2014**. A bayesian method for analyzing lateral gene transfer. *Systematic Biology*, **63**: 409–420. doi:10.1093/sysbio/syu007

Stamatakis, A., **2006**. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**: 2688–2690. doi:10.1093/bioinformatics/btl446

Steglich, W., Klaar, M. & Furtner, W., **1974**. (+)-Mitorubrin derivatives from *Hypoxylon fragiforme*. *Phytochemistry*, **13**: 2874–2875. doi:10.1016/0031-9422(74)80262-1

Stolzer, M., Lai, H., Xu, M., Sathaye, D., Vernot, B., et al., **2012**. Inferring duplications, losses, transfers and incomplete lineage sorting with nonbinary species trees. *Bioinformatics*, **28**: i409–i415. doi:10.1093/bioinformatics/bts386

Szöllősi, G.J., Tannier, E., Daubin, V. & Boussau, B., **2014**. The inference of gene trees with species trees. *Systematic Biology*, **64**: syu048–. doi:10.1093/sysbio/syu048

Vernot, B., Stolzer, M., Goldman, A. & Durand, D., **2008**. Reconciliation with non-binary species trees. *Journal of Computational Biology*, **15**: 981–1006. doi:10.1089/cmb.2008.0092

Wernersson, R. & Pedersen, A.G., **2003**. RevTrans: Multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Research*, **31**: 3537–3539. doi:10.1093/nar/gkg609

Woo, P.C.Y., Lam, C.-W., Tam, E.W.T., Lee, K.-C., Yung, K.K.Y., et al., **2014**. The biosynthetic pathway for a thousand-year-old natural food colorant and citrinin in *Penicillium marneffei*. *Scientific Reports*, **4**: 6728. doi:10.1038/srep06728

Woo, P.C.Y., Lam, C.-W., Tam, E.W.T., Leung, C.K.F., Wong, S.S.Y., et al., **2012**. First discovery of two polyketide synthase genes for mitorubrinic acid and mitorubrinol yellow pigment biosynthesis and implications in virulence of *Penicillium marneffei*. *PLoS neglected tropical diseases*, **6**: e1871. doi:10.1371/journal.pntd.0001871

Wu, Y.-C., Rasmussen, M.D., Bansal, M.S. & Kellis, M., **2013**. TreeFix: Statistically informed gene tree error correction using species trees. *Systematic Biology*, **62**: 110–120. doi:10.5061/dryad.44cb5

Yaegashi, J., Praseuth, M.B., Tyan, S.-W., Sanchez, J.F., Entwistle, R., et al., **2013**. Molecular genetic characterization of the biosynthesis cluster of a prenylated isoindolinone alkaloid aspernidine A in *Aspergillus nidulans*. *Organic Letters*, **15**: 2862–2865. doi:10.1021/ol401187b

Yang, Y., Liu, B., Du, X., Li, P., Liang, B., et al., **2015**. Complete genome sequence and transcriptomics analyses reveal pigment biosynthesis and regulatory mechanisms in an industrial strain, *Monascus purpureus* YY-1. *Scientific Reports*, **5**: 8331. doi:10.1038/srep08331

Zabala, A.O., Xu, W., Chooi, Y.-H. & Tang, Y., **2012**. Characterization of a silent azaphilone gene cluster from *Aspergillus niger* ATCC 1015 reveals a hydroxylation-mediated pyran-ring formation. *Chemistry and Biology*, **19**: 1049–1059. doi:10.1016/j.chembiol.2012.07.004

Chapter 7

Hints of horizontal gene transfer of mitorubrin gene cluster from *Talaromyces* to *Stachybotrys*

This chapter presents initial phylogenetic investigations of a potential example of horizontal gene transfer from *Talaromyces* to *Stachybotrys*. This study was initiated towards the very end of the time as PhD student and further investigations adding support to the results should be conducted before potential publication.

7.1 Introduction

The genetic machinery for biosynthesis of secondary metabolites tends to cluster in filamentous fungi (Keller & Hohn, 1997). Speculations on the reason for the clustering have been many. Among the speculations possibility for co-regulation or potential for rapid exchange of entire gene clusters between species are recurring. In addition, the clustering of SM genes makes it possible for the fungus to eliminate an entire cluster when no ecological pressure is present to retain the gene cluster.

Rapid increase in available fungal genomes has in recent years resulted in several reports of HGT of fungal secondary metabolite cluster. The probably most striking example reported to date is the transfer of the 23 gene large sterigmatocystin (ST) cluster from *Aspergillus* to *Podospora* (Slot & Rokas, 2011). In this case the microsynteny between the ST cluster in *Aspergillus nidulans* and *Podospora anserina* is very high, and the production of ST have also been observed in both species (Matasyoh et al., 2011). Other less clear examples of HGT of SM clusters in filamentous fungi include transfer of the ACE1 gene cluster from *Magnaporthe* to *Aspergillus* (Khaldi et al., 2008), the fumonisin gene cluster between *Fusarium* and *Aspergillus* (Khaldi & Wolfe, 2011) and the bikaverin cluster from *Fusarium* to *Botrytis* (Campbell et al., 2012; Campbell et al., 2013).

In Chapter 5 and 6 I showed the reduction of the *Monascus* pigment (MP) cluster in *Talaromyces atroroseus* IBT11181 by the loss of the non-reducing MP-PKS. This loss has probably been driven by the redundancy in having two PKS genes producing a precursor to be used in two separate azaphilone biosynthetic pathways. The PKS complementing the loss of the MP-PKS showed to be the mitorubrin PKS, MitA. It was shown that the mitorubrin and *Monascus* pigment gene cluster in *Talaromyces* is a result of a gene duplication event and that the *aza* cluster in *Aspergillus niger* is a result of a diverged mitorubrin cluster. During subsequent analysis of the mitorubrin PKS I discovered that *Stachybotrys chartarum* and *Stachybotrys chlorohalonata* have a PKS with high similarity to the mitorubrin PKS. PKS genes with similar high homology to the mitorubrin PKS were also found in *Thielavia terrestris* and *Metarhizium guizhouense*, however none of these have to my knowledge been reported to produce azaphilones. All four species belong to the *Sordariomycetes* and are distant related to *Talaromyces*. Apparently no other species from *Sordariomycetes* contain any closely related PKS genes. This could indicate yet an example of SM genes having undergone HGT between fungal lineages.

The azaphilone class of mitorubrins produced by *Talaromyces* species are also produced by a range of species belonging to *Hypoxylon* including *H. fragiforme* (Steglich et al., 1974). Sir *et al.* (Sir et al., 2015) state that the evolution of the mitorubrin gene cluster within *Hypoxylon* is believed to originate from a common ancestral gene cluster in the *Xylariaceae*, rather than being the result of convergent evolution. However the only *Hypoxylon* genome available in Genbank at the time of the present study (*Hypoxylon* sp. E7406B) does not contain a cluster similar to the mitorubrin cluster in *Talaromyces* species. So, inferring possible relationship between the mitorubrin gene clusters in *Talaromyces* species and mitorubrin producing *Hypoxylon* species is not possible in the current scope of the present study.

In the present study the phylogenetic and evolutionary relationship between the mitorubrin cluster in *Talaromyces* species and the gene clusters containing mitorubrin PKS homologs in *S. chartarum, S. chlorohalonata*, *Th. terrestris* and *M. guizhouense*, are investigated in terms of potential HGT events.

7.2 Materials and Methods

7.2.1 Selection cluster genes for blast

For the construction of individual gene phylogenies, genes with offset in the mitorubrin gene cluster from *T. atroroseus* IBT11181 along with genes up- and downstream of the putative gene cluster boundaries was selected. Specifically, genes ranging from the first gene on the genomic scaffold, UA08_05974, and to the gene UA08_05990, located six genes downstream of the azaphilone NR-PKS gene (*mitA*), were selected (Figure 7.1). To ease the readability for the reader, these genes are named ORF1-17 during the study. Furthermore genes in the *T. marneffei* ATCC18224 mitorburin cluster region, with no homolog in the *T. atroroseus* mitorubrin gene cluster, were also selected. These have been named ORFA-ORFD (Figure 7.1).

All genes were used in BLASTX search against the nr database from NCBI as well as a local database containing the predicted proteomes of *Monascus purpureus* NRRL1596, *Monascus ruber* NRRL1597, *Talaromyces aculeatus* ATCC10409 and *Talaromyces atroroseus* IBT11181. A cut-off e-value of 1e-5 was applied and the top 80 hits were collected for further tree construction. As ORF7 and ORF9 are homologs of each other, they were hits in each other's blast search. Therefore the hits from the two searches were combined resulting in total of 91 hits and used to construct a tree with both ORF7 and ORF9.

7.2.2 Alignments and tree construction

Alignment for each set of proteins were constructed with MAFFT v6.902b using the L-INS-I option (Katoh et al., 2002; Katoh & Toh, 2008). For the PKS set of proteins, i.e. ORF6 and ORF11, the KS domains were extracted and used for the alignment. The alignments were manually inspected and sites with more than 20% missing data were removed with trimAL v1.4 (Capella-Gutiérrez et al., 2009). Bayesian phylogenetic trees where constructed with MrBayes v3.2.1 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003; Ronquist et al., 2012). The MCMC chain was run in duplicate with one cold and three hot chains. The chains were run for 300,000 generations sampling every 100th generation resulting in total 3001 sampled trees. Convergence was evaluated by Tracer (Rambaut et al., 2014) The trees were summarized with MrBayes as 50% majority rule discarding the first 25 % of sampled trees. All consensus trees have been midpoint rooted.

7.3 Results and discussion

7.3.1 Azaphilone gene clusters related to mitorubrin gene cluster

In Figure 7.1 the mitorubrin cluster organisation of *T. atroroseus* IBT11181 and *T. marneffei* ATCC18224 are compared with the azaphilone gene clusters in *Stachybotrys* species. The scaffolds containing the gene cluster in the *Stachybotrys* genomes range from ~28 kb in *S. chlorohalonata* IBT40285 to ~114 kb in *S. chartarum* IBT40293. The rather short scaffolds mean that the actual gene cluster might be spread on several scaffolds. However, the microsynteny of the part of the gene cluster that is present in all four genome sequenced *Stachybotrys*, i.e *S. chartarum* IBT7711, *S. chartarum* IBT40288, *S. chartarum* IBT40293 and *S. chlorohalonata* IBT40285, are identical. Thus, only the gene clusters from *S. chartarum* IBT40293 and *S. chlorohalonata* ÌBT40285 are shown in Figure 7.1. For *S. chartarum* IBT40293 gene calling with AUGUSTUS (Stanke et al., 2004) suggested that two of the genes, S40293_07252 and S40293_07257, should be split in two separate genes. These splits have been implemented in Figure 7.1. Similar splits were predicted in the other *Stachybotrys* gene clusters.

For comparison gene clusters containing PKS genes closely related to the mitorubrin PKS (ORF11) are also shown in Figure 7.1, i.e. the *Th. terrestris* gene cluster, the *M. guizhouense* gene cluster and the *A. niger aza* gene cluster. Initial cluster comparison showed that 5 genes are shared between the mitorubrin cluster and the azaphilone cluster in *Stachybotrys*. These include

the two PKS genes (ORF6 and ORF11) responsible for production of the orsellinic acid and azaphilone part of the mitorubrins from *Talaromyces* species (Woo et al., 2012).

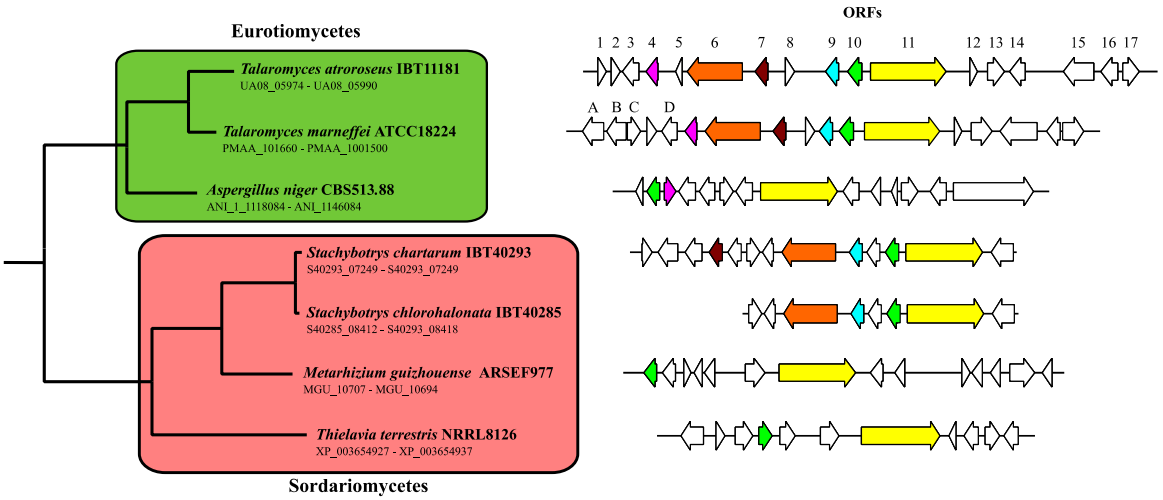


Figure 7.1. Phylogenetic relationship between ascomycetes harbouring gene clusters with PKS gene closely homologous to *T.* atroroseus *mitA gene.* As gene clusters within *Talaromyces* and *Stachybotrys* show identical microsynteny only representative species is shown. The boundaries of the mitorubrin gene cluster in *Talaromyces* are yet unknown, so 5-6 genes away from the two PKS genes were used as analysis boundaries. The genes from *T. atroroseus* are referred to as ORF1-ORF17, and genes from *T. marneffei* with no orthology in *T. atroroseus* are referred to as ORFA-ORFD. Coloured genes identify homologous set of genes across the different species. As the microsynteny of *T. atroroseus* and *T. marneffei* a very similar only genes with homologs in non-*Talaromyces* species have been coloured. For detailed microsynteny of the mitorubrin cluster in *Talaromyces,* see Figure 6.1. For each species the range of locus tags shown is stated. Since the locus tags in *Th. terrestris* is not sequential numbered, the sequential protein accessions is instead stated.

7.3.2 Gene-by-gene phylogenies suggest HGT

To investigate the evolutionary relationship between the mitorubrin cluster in *Talaromyces* and the azaphilone gene cluster in *Stachybotrys* individual gene-by-gene phylogenies with offset in the *T. atroroseus* mitorubrin gene cluster were created at protein level.

The trees with *Stachybotrys* genes in phylogenetic relation to the *T. atroroseus* genes show different topologies and are illustrated in Figure 7.2, Figure 7.3 and appendix Figure E.1-E.8.

The trees of ORF1, ORF10 and ORF11 are presented in Figure 7.2 as they show similar topologies. In the mitorubrin NR-PKS tree (ORF11) the *Stachybotrys* NR-PKS along with the *Th. terrestris* and *M. guizhouense* NR-PKS fall between the clade with *Talaromyces* mitorubrin PKSs and the *A. niger* clade of the NR-PKS from the *aza* cluster. A similar topology is observed in the ORF10 tree. This tree topology is a clear contradiction of the species phylogeny and hence is strongly supportive for a horizontal transfer of these two corresponding genes from *Talaromyces* to a common ancestral *Sordariomycete* of *Stachybotrys*, *Th. terrestris* and *M. guizhouense.*

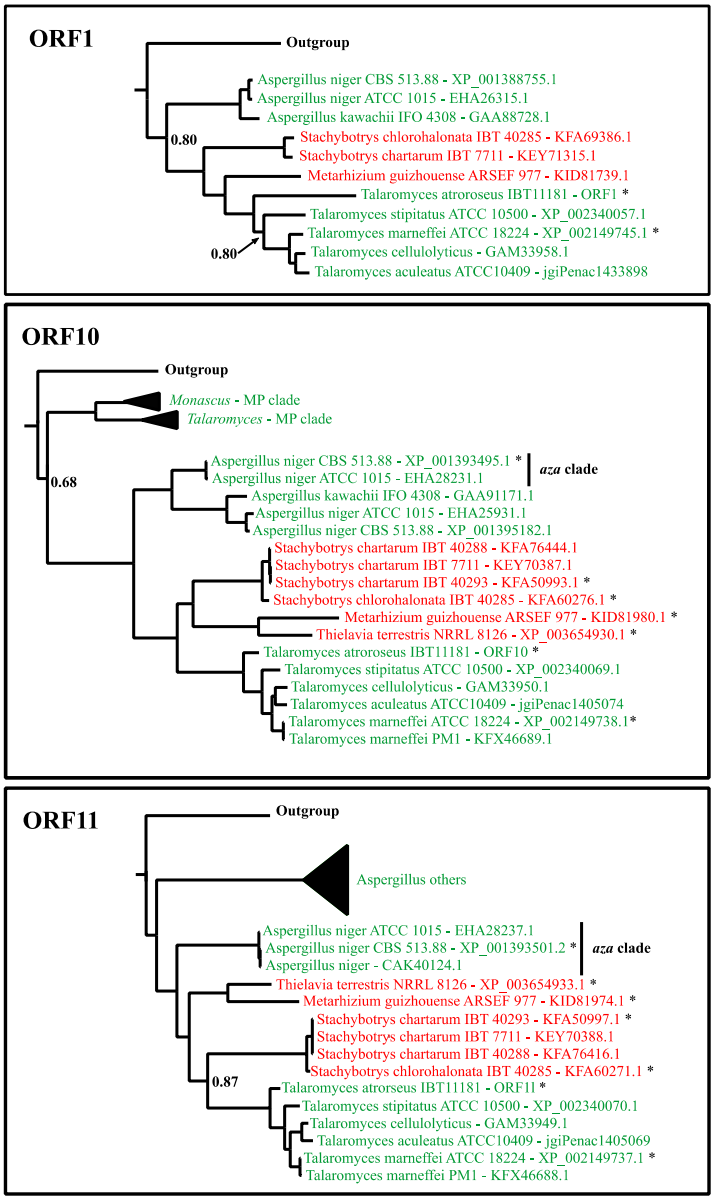


Figure 7.2. Phylogenetic trees of ORF1, ORF10 and ORF11. Trees are summarized as 50% majority rule consensus trees constructed with MrBayes. Each protein has been coloured according to the taxonomical class of the species it is coming from, Green: *Eurotiomycetes*, Red, *Sordariomycetes*. Posterior probabilities are shown on nodes when less than 0.90. Proteins encoded by genes in Figure 7.1 are noted with an asterisk. The full trees can be found in appendix E.

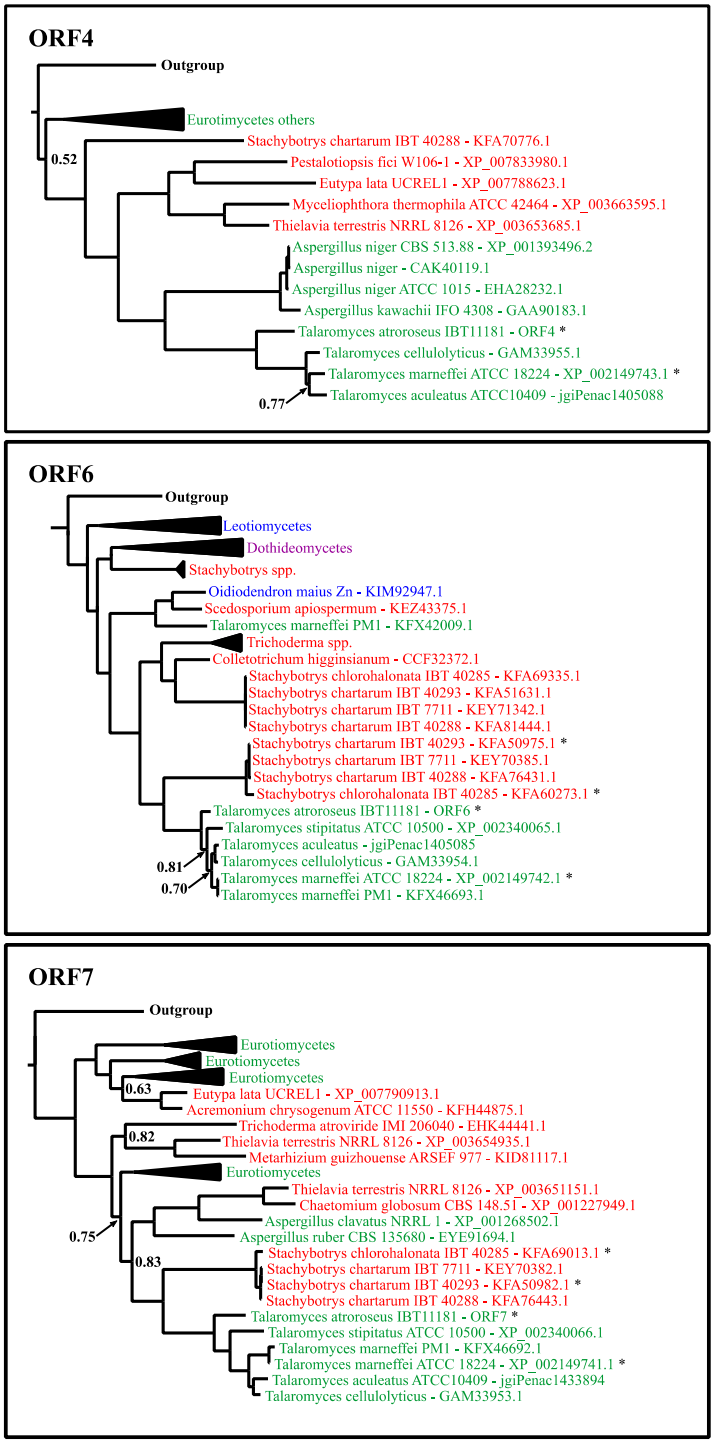


Figure 7.3. Phylogenetic trees of ORF4, ORF6 and ORF7. Trees are summarized as 50% majority rule consensus trees constructed with MrBayes. Each protein has been coloured according to the taxonomical class of the species it is coming from, Green: *Eurotiomycetes*, Red, *Sordariomycetes*, Blue, *Leotiomycetes*. Posterior probabilities are shown on nodes when less than 0.90. Proteins encoded by genes in Figure 7.1 are noted with an asterisk. The full trees can be found in appendix E.

Only two of the four *Stachybotrys* genomes contain a homolog to ORF1, and also *Th. terrestris* miss this homolog. Furthermore the proteins in the *A. niger* clade, as well as the present *Sordariomyctes* proteins are not found in or close to the gene clusters in Figure 7.1. The tree topology however still supports a HGT of ORF1. Hypothesising that the genes clustered in the *Talaromyces* species have been transferred in one event means that the ORF1 gene has undergone differential loss along the different *Sordariomycetes* lineages and has as well been chromosomally relocated. A similar topology is found in the ORFB tree. The ORFB (Appendix E.8) is however not found in the *T. atroroseus* mitorubrin cluster or in *S. chartarum* IBT40293. In the other *Stachybotrys* species the ORFB homolog is also relocated from the putative azaphilone cluster, as is ORF1.

The ORF4 tree (Figure 7.3) only contains a protein from one *Stachybotrys* genome, i.e. *S. chartarum* IBT40288. However, proteins from a couple of other *Sordariomycetes* are present in the tree. These groups together between a clade consisting of the ORF4 from *Talaromyces* species and proteins from *A. niger* and *A. kawachii* and a clade containing other *Eurotiomycetes* proteins. Even though this is indicative of a HGT event, several observations might suggest alternative evolutionary stories. First hint of alternative evolution is the occurrence of different *Sordariomycetes* than those observed in the other phylogenetic trees. Furthermore, the *Sordariomycetes* branch is outside the monophyletic clade with *Talaromyces* and *A. niger* genes. With these observations the tree topology could also be explained by a pattern of duplications followed by losses.

ORF7 and ORF9 are homologs to each other and can be found in the same phylogenetic tree in appendix E (See Figure E.4). Subtrees of Figure E.4 for ORF7 (Figure 7.3) and ORF9, respectively, show similar topologies. Moreover, the topologies are very similar to what was observed for the ORF10 and ORF11 trees discussed above. However, the ORF7 and ORF9 trees do not contain proteins from the *aza* cluster. Instead the *Stachybotrys* clade from the azaphilone cluster falls between the *Talaromyces* gene cluster and other *Aspergillus* and *Eurotiomycetes* proteins. This supports a HGT as above, but with the ORF7 and ORF9 homologs lost from the *A. niger aza* cluster. On the other hand, with other *Sordariomycetes* proteins also appearing within the *Eurotiomycetes* clade, an evolutionary route with ancient duplications followed by massive loss of the corresponding genes cannot be entirely excluded as an explanation of the resulting tree topology.

The phylogenetic tree of the putative orsellinic acid PKS (ORF6, Figure 7.3) show a close relation between the PKS found in the *Talaromyces* and the *Stachybotrys* gene clusters, as they

group into sister groups. However, these clades fall within a clade mainly encompassing *Sordariomycetes*. This suggests that the ORF6 gene has been transferred in the opposite direction than discussed for the other topologies.

In the ORF12 tree (Figure E.7) a clade of *Stachybotrys* proteins appear as sister group to the mitorubrin cluster proteins from *Talaromyces*. However the tree is poorly resolved as most of the other clades appear in one big polytomy, making it impossible to infer whether this tree follow the species topology or not.

The ORF3 and ORF5 trees contain one big polytomy making it impossible to infer any evolutionary relationships. In none of these trees, proteins encoded by genes in close proximity to any of the *Sordariomycetes* azaphilone PKS genes in Figure 7.1 are present (data not shown). The ORF14 tree contains a clade of *Stachybotrys* proteins and the genes are not associated with the putative azaphilone cluster in Figure 7.1. These proteins cluster together with other *Sordariomycetes* proteins from *Fusarium*, *Podospora* and *Magnaporthe* species. The *Talaromyces* proteins in all these mentioned trees do not contain any close homologs with potential HGT involved evolutionary relationship and are hence not shown. The resulting gene trees ORF2, ORF8, ORF13, ORF15, ORF16, ORF17, ORFA, ORFC and ORFD do not contain any proteins from *Stachybotrys*, *Metharhizium* or *Thielavia* in relation to the *T. atroroseus* genes used as query. In addition, the *Talaromyces* proteins group together with other *Eurotiomycetes* proteins following the corresponding species phylogeny. Hence, these trees are neither shown here nor in Appendix.

In summary no evolutionary hypothesis covers all the gene-by-gene phylogenies. However, ORF10 and ORF11, supported by several of the other phylogenies, strongly suggest that an ancestral mitorubrin cluster has undergone HGT from *Talaromyces* to an ancestral *Sordariomycete*. This cluster has as minimum consisted of ORF10 and ORF11 and potentially ORF1, ORF4, ORF7 and ORF9. The transfer of the cluster has occurred after the speciation of *Aspergillus* and *Talaromyces*, since the *Stachybotrys* proteins cluster between the clades with the mitorubrin gene cluster proteins from *Talaromyces* and the *A. niger aza* cluster proteins.

The inferred HGT could potential have happened before speciation of the *Stachybotrys*, *Thielavia* and *Metarhizium* lineages and the transferred cluster have then undergone massive reduction in the latter two only retaining the ORF10 and ORF11 homologs. Similar reduction of transferred SM clusters have been observed in the bikaverin cluster in *Botrytis* (Campbell et al., 2012).

As stated by Sir *et al.* (Sir et al., 2015) the evolution of the mitorubrin and similar azaphilone clusters in *Ascomycetes* is very unlikely to be due to convergent evolution. The phylogenies presented in this study reveals that the mitorubrin cluster might have a much more complex evolutionary history than just vertical inheritance and cluster loss in non-mitorubrin producing genera and species. In the future, when these *Hypoxylon* species have been genome sequenced and made available for community research, interesting phylogenetic analyses should include data from mitorubrin-producing *Hypoxylon* species to enlighten the evolution of the mitorubin gene cluster.

7.4 References

Campbell, M.A., Rokas, A. & Slot, J.C., **2012**. Horizontal transfer and death of a fungal secondary metabolic gene cluster. *Genome Biology and Evolution*, **4**: 289–293. doi:10.1093/gbe/evs011

Campbell, M.A., Staats, M., van Kan, J.A.L., Rokas, A. & Slot, J.C., **2013**. Repeated loss of an anciently horizontally transferred gene cluster in *Botrytis*. *Mycologia*, **105**: 1126–34. doi:10.3852/12-390

Capella-Gutiérrez, S., Silla-Martínez, J.M. & Gabaldón, T., **2009**. TrimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, **25**: 1972–1973. doi:10.1093/bioinformatics/btp348

Huelsenbeck, J.P. & Ronquist, F., **2001**. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**: 754–755. doi:10.1093/bioinformatics/17.8.754

Katoh, K., Misawa, K., Kuma, K. & Miyata, T., **2002**. MAFFT : a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acid Research*, **30**: 3059–3066.

Katoh, K. & Toh, H., **2008**. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, **9**: 286–98. doi:10.1093/bib/bbn013

Keller, N.P. & Hohn, T.M., **1997**. Metabolic pathway gene clusters in filamentous fungi. *Fungal Genetics and Biology*, **21**: 17–29. doi:10.1006/fgbi.1997.0970

Khaldi, N., Collemare, J., Lebrun, M.-H. & Wolfe, K.H., **2008**. Evidence for horizontal transfer of a secondary metabolite gene cluster between fungi. *Genome Biology*, **9**: R18. doi:10.1186/gb-2008-9-1-r18

Khaldi, N. & Wolfe, K.H., **2011**. Evolutionary origins of the fumonisin secondary metabolite gene cluster in *Fusarium verticillioides* and *Aspergillus niger*. *International Journal of Evolutionary Biology*, **2011**: 423821. doi:10.4061/2011/423821

Matasyoh, J.C., Dittrich, B., Schueffler, A. & Laatsch, H., **2011**. Larvicidal activity of metabolites from the endophytic *Podospora sp.* against the malaria vector *Anopheles gambiae*. *Parasitology Research*, **108**: 561–566. doi:10.1007/s00436-010-2098-1

Rambaut, A., Suchard, M., Xie, D. & Drummond, A.J., **2014**. Tracer v1.6. Available from http://beast.bio.ed.ac.uk/Tracer

Ronquist, F. & Huelsenbeck, J.P., **2003**. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**: 1572–1574. doi:10.1093/bioinformatics/btg180

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., et al., **2012**. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**: 539–542. doi:10.1093/sysbio/sys029

Sir, E.B., Kuhnert, E., Surup, F., Hyde, K.D. & Stadler, M., **2015**. Discovery of new mitorubrin derivatives from *Hypoxylon fulvo-sulphureum* sp. nov. (Ascomycota, Xylariales). *Mycological Progress*, **14**: 28. doi:10.1007/s11557-015-1043-1

Slot, J.C. & Rokas, A., **2011**. Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. *Current Biology*, **21**: 134–139. doi:10.1016/j.cub.2010.12.020

Stanke, M., Steinkamp, R., Waack, S. & Morgenstern, B., **2004**. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Research*, **32**: W309–W312. doi:10.1093/nar/gkh379

Steglich, W., Klaar, M. & Furtner, W., **1974**. (+)-Mitorubrin derivatives from *Hypoxylon fragiforme*. *Phytochemistry*, **13**: 2874–2875. doi:10.1016/0031-9422(74)80262-1

Woo, P.C.Y., Lam, C.-W., Tam, E.W.T., Leung, C.K.F., Wong, S.S.Y., et al., **2012**. First discovery of two polyketide synthase genes for mitorubrinic acid and mitorubrinol yellow pigment biosynthesis and implications in virulence of *Penicillium marneffei*. *PLoS Neglected Tropical Diseases*, **6**: e1871. doi:10.1371/journal.pntd.0001871

Chapter 8
Concluding remarks and perspectives

Modern research within fungal natural products requires researchers with ability to adapt too many different scientific areas within biotechnology, mycology, analytical chemistry and bioinformatics. A modern researcher should apart from being a specialist in selected fields, also have the ability to bridge between very different scientific fields.

The importance of bridging between different fields has been paramount in finishing the present study. A multitude of scientific areas and techniques have been used. These have involved analytical methods for metabolite profiling of *T. atroroseus* as suitable pigment producer, designing NGS experiments as well as assembling and annotating genomes, molecular biology techniques, as well as inferring evolutionary relationship between azaphilone gene clusters in fungal species.

In the present PhD thesis important steps towards optimising and using *Talaromyces atroroseus* as cell factory for production of *Monascus* pigments have been taken. Importantly, the reproducibility of controlled reactor-based fermentation of *T. atroroseus* shown in Chapter 4 is a necessary basis for the continued development of *T. atroroseus* production conditions for red pigment production. However several matters should be of consideration in relation to use *T. atroroseus* for industrial production of *Monascus* pigments. The observed growth rate of *T. atroroseus* IBT11181 is significant lower than the growth rate from other industrial fungal strains. Since food colorants is a low value product, it is important to consider optimisation of growth conditions to get a higher growth rate in order to make the pigment production a profitable process. This optimisation of fermentation conditions should however take place without potential loss of pigment production. This research track is already ongoing at Section for Eukaryotic Biotechnology by PhD student Gerit Nymschefsky. Furthermore, genetic

improvements of the IBT11181 strain should be considered for better physiological performance during fermentation.

Apart from strain improvements for better physiological performance under fermentation, genes directly involved in the biosynthesis of *Monascus* pigments are obvious targets for strain improvements. Obvious targets to enhance the red pigment production during sub-merged fermentation are among other over-expression of the transcriptional activator in the MP cluster. The MP cluster also contains a putative MFS transporter which might work by facilitating the secretion of the produced MPs. This is also an obvious target for over-expression, potentially creating a strain with enhanced secretion capabilities. The biosynthesis of MPs might be limited by the concentration of produced MPs in the mycelia, and hence an increased secretion will potentially also lead to a higher pigment yield.

The genome sequence provided to the research community as part of this study will be a valuable tool for potential strain improvement of *T. atroroseus* IBT11181. However, for strain improvements as well as further study of the azaphilone biosynthesis, enhancement of the molecular toolbox will be of great importance. Especially, the availability of a non-homolog end-joining deficient strain will be of excellent use to improve success rate in genetic transformations. Also development or adaptation of a recyclable marker system, from e.g. *A. nidulans*, should be on the wish list. The CRISPR/Cas system for effectively genetic engineering, has recently been adapted and shown to work in wild type filamentous fungi such as *Trichoderma reseii* (Liu et al., 2015)*, Magnaporthe oryzae* (Arazoef et al., 2015) and a selection of *Aspergillus* species (Nødvig et al., 2015). Adaptation and implementation of these tools in *T. atroroseus* would be a logical and important next step.

One concern in the industrial production of *Monascus* pigments is the co-production of the mycotoxin citrinin. Production of citrinin has generally been associated with the *Monascus* genus as well as some *Aspergillus* and *Penicillium* species such as *Aspergillus terreus, Penicillium citrinum* and *P. expansum* (Frisvad et al., 2007). However recently the *pks3 Monascus* pigment PKS in *T. marneffei* was linked to also be responsible for citrinin production (Woo et al., 2014). This was the first observation of citrinin production within the *Talaromyces* genus and the finding indicates that other *Monascus* pigment producing *Talaromyces* strains may also have the machinery to produce citrinin. The production of citrinin will limit the potential use of *Talaromyces* species as alternative cell factories for red pigment production. As citrinin has so far only been observed once within *Talaromyces* more research is necessary to address the

citrinin-production ability of other *Talaromyces*. *T. atroroseus* IBT11181 do however not contain the *Monascus* PKS making it still an attractive choice for further studying as red pigment producer.

The distribution of secondary metabolite gene clusters across fungal genomes is truly interesting. This sometimes scattered nature of some SM gene clusters is valuable in terms of understanding the evolutionary trajectories of gene clusters in fungi. However, the mechanisms involved in potential HGT events between filamentous fungi are still unknown (Wisecaver & Rokas, 2015). In the coming years the number of available fungal genomes will almost explode. This will make the phylogenetic analysis of evolution of SM clusters much more comprehensive and robust and provide an even better understanding of evolution of SM clusters in fungi.

The wide distribution of SM gene clusters also provides important insights into the complex biosynthesis of fungal SMs. The continued increase in the understanding of these biosynthetic pathways and gene functions, provide a platform for the future shuffling of genes and creation of complete novel gene clusters, with potential of producing the next medicinal drug.

8.1 References

Arazoef, T., Miyoshi, K., Yamato, T., Ogawa, T., Ohsato, S., et al., **2015**. Tailor-made CRISPR/Cas system for highly efficient targeted gene replacement in the rice blast fungus. *Biotechnology and Bioengineering*, **2015**: n/a–n/a. doi:10.1002/bit.25662

Frisvad, J.C., Thrane, U. & Samson, R.A., **2007**. Mycotoxin producers. In J. Dijksterhuis & R. A. Samson, eds. *Food Mycology: A Multifaceted Approach to Fungi and Food*. CRC Press, pp. 135–160.

Liu, R., Chen, L., Jiang, Y., Zhou, Z. & Zou, G., **2015**. Efficient genome editing in filamentous fungus *Trichoderma reesei* using the CRISPR/Cas9 system. *Cell Discovery*, **1**: 15007. doi:10.1038/celldisc.2015.7

Nødvig, C.S., Nielsen, J.B., Kogle, M.E. & Mortensen, U.H., **2015**. A CRISPR-Cas9 system for genetic engineering of filamentous fungi. *PLOS ONE*, **10**: e0133085. doi:10.1371/journal.pone.0133085

Wisecaver, J.H. & Rokas, A., **2015**. Fungal metabolic gene clusters: caravans traveling across genomes and environments. *Frontiers in Microbiology*, **6**: 1–11. doi:10.3389/fmicb.2015.00161

Woo, P.C.Y., Lam, C.-W., Tam, E.W.T., Lee, K.-C., Yung, K.K.Y., et al., **2014**. The biosynthetic pathway for a thousand-year-old natural food colorant and citrinin in *Penicillium marneffei*. *Scientific Reports*, **4**: 6728. doi:10.1038/srep06728

Technical University of Denmark

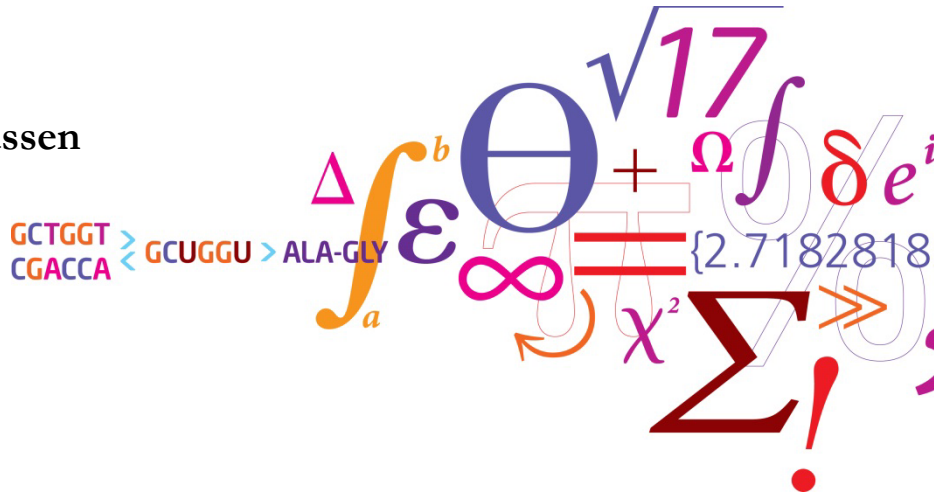
APPENDICES TO

TALAROMYCES ATROROSEUS

GENOME SEQUENCING, *MONASCUS* PIGMENTS AND AZAPHILONE GENE CLUSTER EVOLUTION

Kasper Bøwig Rasmussen

PhD Thesis
August 2015



Content of Appendices

Appendix A

Supporting information to Chapter 3

This appendix contains the published manuscript with the title "*Talaromyces atroroseus*, a new species efficiently producing industrially relevant red pigments" along with its supplementary material in the form of a table of compounds searched for by UHPLC-HR-MS.

A.1 PLOS ONE publication

OPEN ACCESS Freely available online PLOS | ONE

Talaromyces atroroseus, a New Species Efficiently Producing Industrially Relevant Red Pigments

Jens C. Frisvad[1*], Neriman Yilmaz[2,3], Ulf Thrane[1], Kasper Bøwig Rasmussen[1], Jos Houbraken[2], Robert A. Samson[2]

1 Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark, **2** CBS–KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands, **3** Department of Biology, Utrecht University, Utrecht, The Netherlands

Abstract

Some species of *Talaromyces* secrete large amounts of red pigments. Literature has linked this character to species such as *Talaromyces purpurogenus, T. albobiverticillius, T. marneffei*, and *T. minioluteus* often under earlier *Penicillium* names. Isolates identified as *T. purpurogenus* have been reported to be interesting industrially and they can produce extracellular enzymes and red pigments, but they can also produce mycotoxins such as rubratoxin A and B and luteoskyrin. Production of mycotoxins limits the use of isolates of a particular species in biotechnology. *Talaromyces atroroseus* sp. nov., described in this study, produces the azaphilone biosynthetic families mitorubrins and *Monascus* pigments without any production of mycotoxins. Within the red pigment producing clade, *T. atroroseus* resolved in a distinct clade separate from all the other species in multigene phylogenies (ITS, β-tubulin and *RPB1*), which confirm its unique nature. *Talaromyces atroroseus* resembles *T. purpurogenus* and *T. albobiverticillius* in producing red diffusible pigments, but differs from the latter two species by the production of glauconic acid, purpuride and ZG–1494α and by the dull to dark green, thick walled ellipsoidal conidia produced. The type strain of *Talaromyces atroroseus* is CBS 133442

Citation: Frisvad JC, Yilmaz N, Thrane U, Rasmussen KB, Houbraken J, et al. (2013) *Talaromyces atroroseus*, a New Species Efficiently Producing Industrially Relevant Red Pigments. PLoS ONE 8(12): e84102. doi:10.1371/journal.pone.0084102

Editor: Scott E. Baker, Pacific Northwest National Laboratory, United States of America

Received July 11, 2013; **Accepted** November 8, 2013; **Published** December 19, 2013

Funding: Part of this work was supported by the Danish Research Agency for Technology and Production Grant 09-064967 and an equipment grant from Agilent Technologies. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: A commercial funder (Agilent Technologies) has provided some instruments via an Agilent Thought Leader Award to Jens C. Frisvad. This commercial funder (along with any other relevant declarations relating to employment, consultancy, patents, products in development or marketed products etc.) does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: jcf@bio.dtu.dk

Introduction

Monascus species are known to produce six major azaphilone pigments being the yellow monascin and ankaflavin; the orange monascorubrin and rubropunctatin and the red monascorubramine and rubropunctamine, in addition to more than 20 related pigments [1,2]. Another azaphilone series of yellow pigments is even more widespread in *Talaromyces*, i.e. the mitorubrins [3–5]. The red pigment producer *Monascus purpureus* has been used primarily in Southern China, Japan and Southeast Asia for making red rice wine, red soybean cheese and Anka (red rice) [6]. A problem is that some samples of *Monascus*–fermented rice have been found to contain the mycotoxin citrinin [7], but also that *Monascus* isolates also often produce mevinolin, a drug that is also unwanted in foods [2]. The production of such mycotoxins and drugs limits the use of *Monascus* for industrial purposes, but since citrinin has not been found in any *Talaromyces* species, the latter may be a good alternative for red pigment production.

Studies have shown that polyketide azaphilone *Monascus* red pigments and/or their amino acid derivatives are naturally produced by *Talaromyces aculeatus*, *T. pinophilus*, *T. purpurogenus* and *T. funiculosus* [8,9]. *Talaromyces amestolkiae*, *T. ruber* and *T. stollii* also produce azaphilone polyketides, as recently described by Yilmaz et al. [10], but in those three species the pigment are not diffusing into the growth medium. *Talaromyces amestolkiae* and *T. stollii* were isolated from immuno-compromised patients and are potential human pathogens, while *T. purpurogenus* produces mycotoxins such as rubratoxins A and B, rugulovasins, and luteoskyrin [10]. These factors limit the use of these species for biotechnological production of azaphilone pigments.

In the current study we describe a new *Talaromyces* species, *T. atroroseus*, which secretes large amounts of *Monascus* red pigments, without the production of any known mycotoxins.

Materials and Methods

Strains

Cultures were obtained from the CBS-KNAW Fungal Biodiversity Centre culture collection, Utrecht, the Netherlands. Fresh isolates deposited in the working collection of the Department of Applied and Industrial Mycology (DTO) housed at CBS, and strains from the IBT collection at DTU Systems Biology in Kongens Lyngby, Denmark were also included in this study. Strains are listed in Table 1. KAS strain numbers are from the fungal collection of Keith A. Seifert, Ottawa, Canada.

Morphological analysis

Macroscopic characters were studied on agar media Czapek-Dox yeast autolysate agar (CYA), CYA supplemented with 5 % NaCl (CYAS), yeast extract sucrose agar (YES), creatine sucrose agar (CREA), dichloran 18 % glycerol agar (DG18), oatmeal agar (OA) and malt extract agar (Oxoid) (MEA). The isolates were also tested on CYA at 37 °C and on Blakeslee malt extract agar (MEA2). All media were prepared as described by Samson et al. [11]. The strains were inoculated in three points onto media in 90-mm Petri dishes and incubated for 7 d at 25 °C in darkness. After incubation, the colony diameters on the various agar media were measured. Colonies were photographed with a Canon EOS 400D. Species were characterized microscopically by preparing slides from MEA. Lactic acid was used as mounting fluid. Specimens were examined using a Zeiss AxioSkop2 plus microscope.

DNA extraction, PCR amplification and sequencing

Strains were grown for 7 to 14 d on MEA prior to DNA extraction. DNA was extracted using the Ultraclean™ Microbial DNA isolation Kit (MoBio, Solana Beach, U.S.A.). The extracted DNA was stored at -20 °C. The ITS regions, regions of the β-tubulin and *RPB1* genes were amplified and sequenced according to methods previously described [12–15].

Data analysis

Sequence contigs were assembled using Seqman from DNAStar Inc. Newly generated ITS, β-tubulin and *RPB1* sequences were included in a data set obtained from the Samson et al. [15] study. Data sets were aligned using Muscle software within MEGA5 [16]. Neighbour–joining analysis on the individual data sets was performed in MEGA5 and confidence in nodes determined using bootstrap analysis with 1000 replicates. *Talaromyces galapagensis* (CBS 751.74^T) was selected as a suitable out-group in all the phylogenies. The newly generated sequences were deposited in GenBank (accession numbers, see Table 1 and Figures 1–3).

Extrolites

Cultures grown on CYA and YES for 7 d at 25 °C were used for extrolite extractions. Extracts were analysed by HPLC using alkylphenone retention indices and diode array UV–VIS detection as described by 17–19, using three 6 mm agar plugs. Standards of extrolites from the collection at DTU Systems

Table 1. Strains used in this study of *Talaromyces atroroseus* and related species.

CBS No.	Other Collection No.	Species	Information and Origin
206.89	IFO 6580, IBT 3960; DTO 41F4	*T. albobiverticillius*	Unknown, Japan
238.95	IBT 11181, CBS 123796	*T. atroroseus*	Red sweet bell pepper, Kgs. Lyngby, Denmark
234.60	DTO 37A4	*T. atroroseus*	Unknown, Germany
257.37	DTO 37A3	*T. atroroseus*	Ex air in nitrite factory, Germany
313.63	DTO 41G2	*T. albobiverticillius*	*Vitis vinifera* fruit, South Africa
364.48	ATCC 9777, IMI 040037, NRRL 1061, QM 6760, DTO 178A3, IBT 4458, IBT 11180	*T. atroroseus*	Unknown, Darien, Manchuria, China
391.96	DTO 41G8	*T. atroroseus*	Unknown, Tanzania
113139	IBT 3967, NRRL 1147, DTO 177I2	*T. atroroseus*	Unknown, USA
113167	DTO 39I2, DTO 39I3	*T. albobiverticillius*	Unknown, unknown
113168	IBT 31347, DTO 39H9, DTO 177I9	*T. albobiverticillius*	Sputum of patient, male, Copenhagen, Denmark
113153	IBT 3458, NRRL 1136, DTO 37A7	*T. atroroseus*	Ex mixed culture, Arlington Farm, Virginia USA
124294	IBT 23082	*T. atroroseus*	Tropical rainforest, Peru
133440	BCRC 34774, DTO 166E5, IBT 31667	Type of *T. albobiverticillius*	Decaying leaves of a broad–leaved tree, Taiwan
133441	BCRC 34775, DTO 166E6, IBT 31668	*T. albobiverticillius*	Decaying leaves of a broad–leaved tree. Taiwan
133442	KAS 3778, DTO 178A4, IBT 32470	Type of *T. atroroseus*	House dust, South Africa
133443	IBT 29388, DTO 189D4	*T. atroroseus*	Contamination in petri dish, Lyngby, Denmark
133444	IMI 163167, IBT 23702, DTO 189C2	*T. albobiverticillius*	*Punica granata*, unknown
133447	DTO 81I2	*T. atroroseus*	Swab sample from cheese warehouse, the Netherlands
133448	DTO 157G5	*T. albobiverticillius*	Pomegranate, Turkey
133449	IBT 29464, DTO 189D5	*T. atroroseus*	Mouse dung, Høve Strand, Denmark
133450	FRR 75, IBT 4454, DTO 188I9	*T. atroroseus*	Soil Murrumbidgee irrigation Area, New South Wales, Australia
133452	NRRL 2120, IBT 3547, DTO 193H9	*T. albobiverticillius*	Cotton duck, Panama
113154	R.B., IMI 090178, NRRL 1214, IBT 3645,	*T. atroroseus*	"Parasite" in *Aspergillus niger* culture,

Table 1 (continued).

CBS No.	Other Collection No.	Species	Information and Origin
	IBT 4428, CBS 127571		Kansas City, Missouri, USA
	TA85S-28-H2, AZ ,IAM 15392, JCM 23216, IBT 32650	*T. atroroseus*	Soil, Thailand
	IBT 20955	*T. atroroseus*	Air root in white mangrove,
			Can de Aruca, Paria Bay, Venezuela
	IBT 4466	*T. albobiverticillius*	*Punica granata*, imported to Denmark

doi: 10.1371/journal.pone.0084102.t001

Biology (Denmark) were used to compare the extrolites from the species under study [18].

The extrolite extractions from *T. atroroseus* CBS 133450, CBS 113154 and CBS 123796 were also analysed by ultra high performance liquid chromatography high-resolution mass spectrometry (UHPLC-HRMS). Liquid chromatography was performed on an Agilent 1290 Infinity LC system with a DAD-detector coupled to an Agilent 6550 iFunnel Q-TOF with an electrospray ionization source. The separation was performed on a 2.1 x 250 mm, 2.7 µm Poroshell 120 Phenyl-Hexyl column (Agilent) at 60 °C with a water-acetonitrile gradient (both with 20 mM formic acid) going from 10 % (vol/vol) to 100 % acetonitrile in 15 min followed by 2.5 min with 100 % acetonitrile and then returning to the start conditions for 2.5 min for equilibration before next sample. All time the flow rate was kept at 0.35 mL/min. HRMS was performed in ESI^+ and extrolites were identified with targeted search on accurate mass of $[M+H]^+$ and $[M+Na]^+$ using Agilent MassHunter Qualitative Analysis B.06.00 software and a database of potential extrolites in *T. atroroseus* with support from UV-VIS spectra. The list of compounds searched for including the extrolite standards can be found in Table S1.

Nomenclature

1. The electronic version of this article in Portable Document Format (PDF) in a work with an ISSN or ISBN will represent a published work according to the International Code of Nomenclature of algae, fungi, and plants, and hence the new names contained in the electronic publication of a PLOS ONE article are effectively published under that Code from the electronic edition alone, so there is no longer any need to provide printed copies. In addition, new names contained in this work have been submitted to MycoBank from where they will be made available to the Global Name Index. The unique MycoBank number can be resolved and the associated information viewed through any standard web browser by appending the MycoBank number contained in this publication to the prefix http://www.mycobank.org/MB. The online version of this work is archived and available from the following digital repositories. PubMed Central, LOCKSS.

2. Repository of *Talaromyces atroroseus* Yilmaz, Frisvad, Houbraken & Samson 2013 sp. nov. [urn:lsid:mycobank.org:804901]

Results and Discussion

The relationship between the *Talaromyces atroroseus* sp. nov. and its close relatives were studied using multigene phylogenies, bason on ITS, *RPB1* and β-tubulin sequences. The aligned datasets were 482, 888 and 374 bp long, respectively. The new species resolved in a clade together with other red pigment producing species such as *T. albobiverticillius*, and *T. minioluteus*. *Talaromyces purpurogenus* resolved in a distantly related clade (Figures 1–3). Within the red pigment producing clade, *T. atroroseus* resolved in a distinct clade separate from all the other species in all three phylogenies, confirming its unique nature.

Historically red pigment production caused a lot of confusion and resulted in numerous misidentifications in literature. This is especially true for *Talaromyces purpurogenus*, *T. ruber*, *Penicillium sanguineum* and *P. crateriforme*. *Penicillium purpurogenum* and *P. rubrum* were described by Stoll [20]. In their monograph Raper and Thom [21] also described *P. purpurogenum* and *P. rubrum*. No type material was available for *P. rubrum* therefore Raper and Thom [21] used two strains to describe *P. rubrum*, NRRL 1062 (= CBS 370.48) and NRRL 2120 (= CBS 133452). Pitt [22] synonymized *P. rubrum*, *P. crateriforme* and *P. sanguineum* with *P. purpurogenum*. The issues in the *T. purpurogenus* complex were clarified by Yilmaz et al. [10] who synonymized *Penicillium crateriforme* and *P. sanguineum* with *T. purpurogenus* and they described *T. ruber* as a distinct species. NRRL 1062 remained as *T. ruber* but NRRL 2120 (= CBS 133452) is a different species than *T. ruber*. Our results showed that NRRL 2120 is *T. albobiverticillius*. Raper and Thom [21] based the *Penicillium purpurogenum* description on NRRL 1061 (= CBS 364.48). However our results show that NRRL 1061 is a typical *T. atroroseus* strain.

Both *Talaromyces purpurogenus* and *T. atroroseus* are common in soil, indoor environments, and fruits. *Talaromyces atroroseus* resembles *T. purpurogenus* and *T. albobiverticillius* in producing red diffusible pigments, but differs from the latter two species by the production of glauconic acid, purpuride and ZG–1494α (Table 2 and Figure 4) and by the dull to dark green thick walled ellipsoidal conidia produced. Barton et al. [26,27] and Barton and Sutherland [28] reported glauconic acid from *P. purpurogenum* IMI 090178, which in the present study has been re-identified as *T. atroroseus*, while ZG–1494α was reported from *P. rubrum* CBS 238.95 [36], which is also a typical *T. atroroseus*. *Talaromyces atroroseus*, *T. purpurogenus* and *T. albobiverticillius* differ from *T. ruber*, *T. amestolkiae* and *T. stollii* by their production of red diffusible pigment. In Table 3 many red pigment producers identified as *Penicillium* species are listed, that may either be *T. purpurogenus, T. ruber, T. albobiverticillius* or *T. atroroseus*. The strains listed in Table 3 were not available for us, so their exact identity cannot be verified.

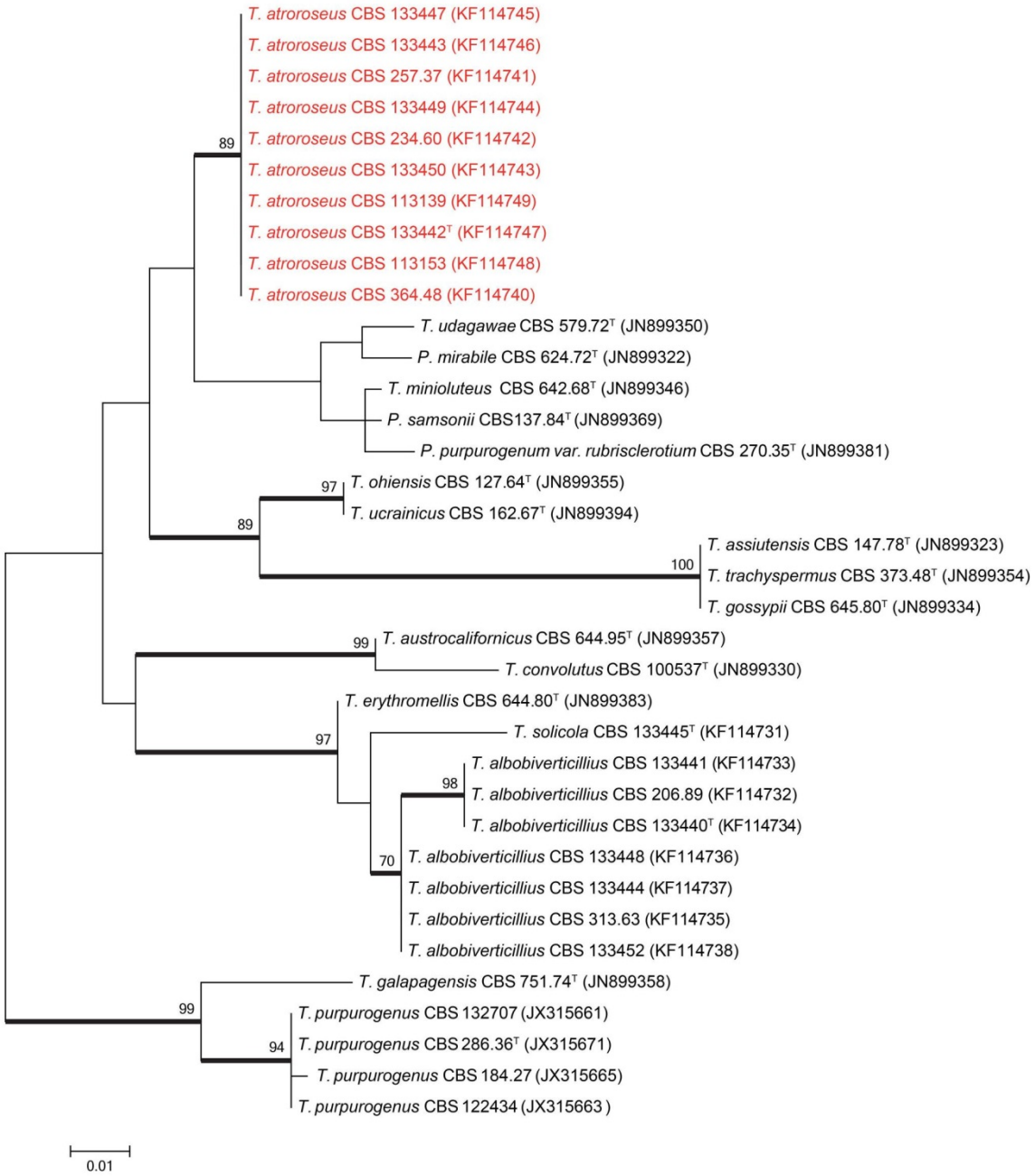


Figure 1. Maximum likelihood tree comparing the ITS gene region of *Talaromyces* species closely related to *T. atroroseus*. *Talaromyces galapagensis* and *T. purpurogenus* were used as outgroup. Support in nodes is indicated above thick branches and is represented by bootstrap values higher than 70%. GenBank accession numbers are given between brackets, (T = ex-type). Red coloured names indicate *T. atroroseus* strains.

doi: 10.1371/journal.pone.0084102.g001

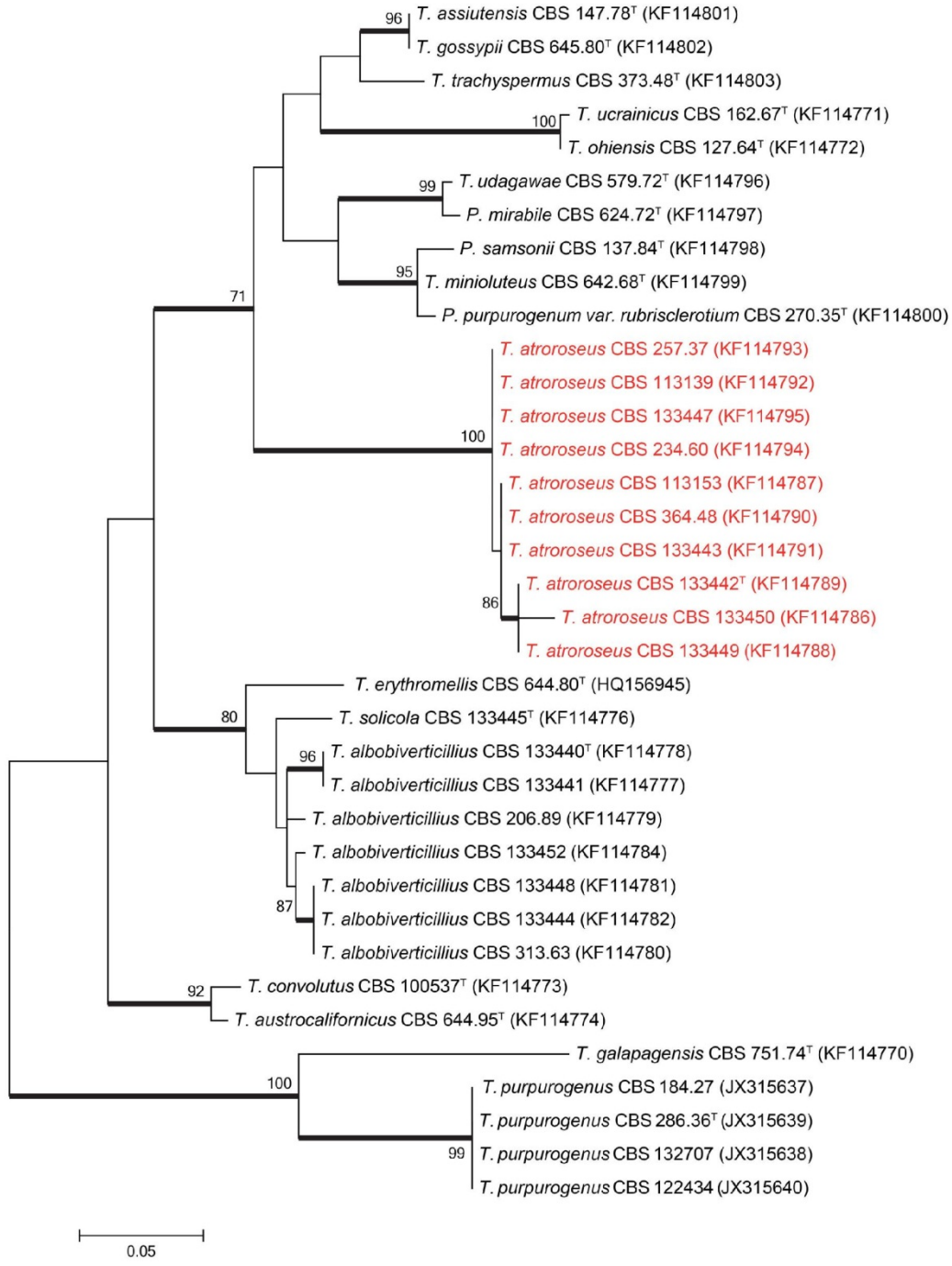


Figure 2. Maximum likelihood tree comparing the β–tubulin gene region of *Talaromyces* species closely related to *T. atroroseus*. *Talaromyces galapagensis* and *T. purpurogenus* were used as outgroup. Support in nodes is indicated above thick branches and is represented by bootstrap values higher than 70%. GenBank accession numbers are given between brackets, (T = ex-type). Red coloured names indicate *T. atroroseus* strains.

doi: 10.1371/journal.pone.0084102.g002

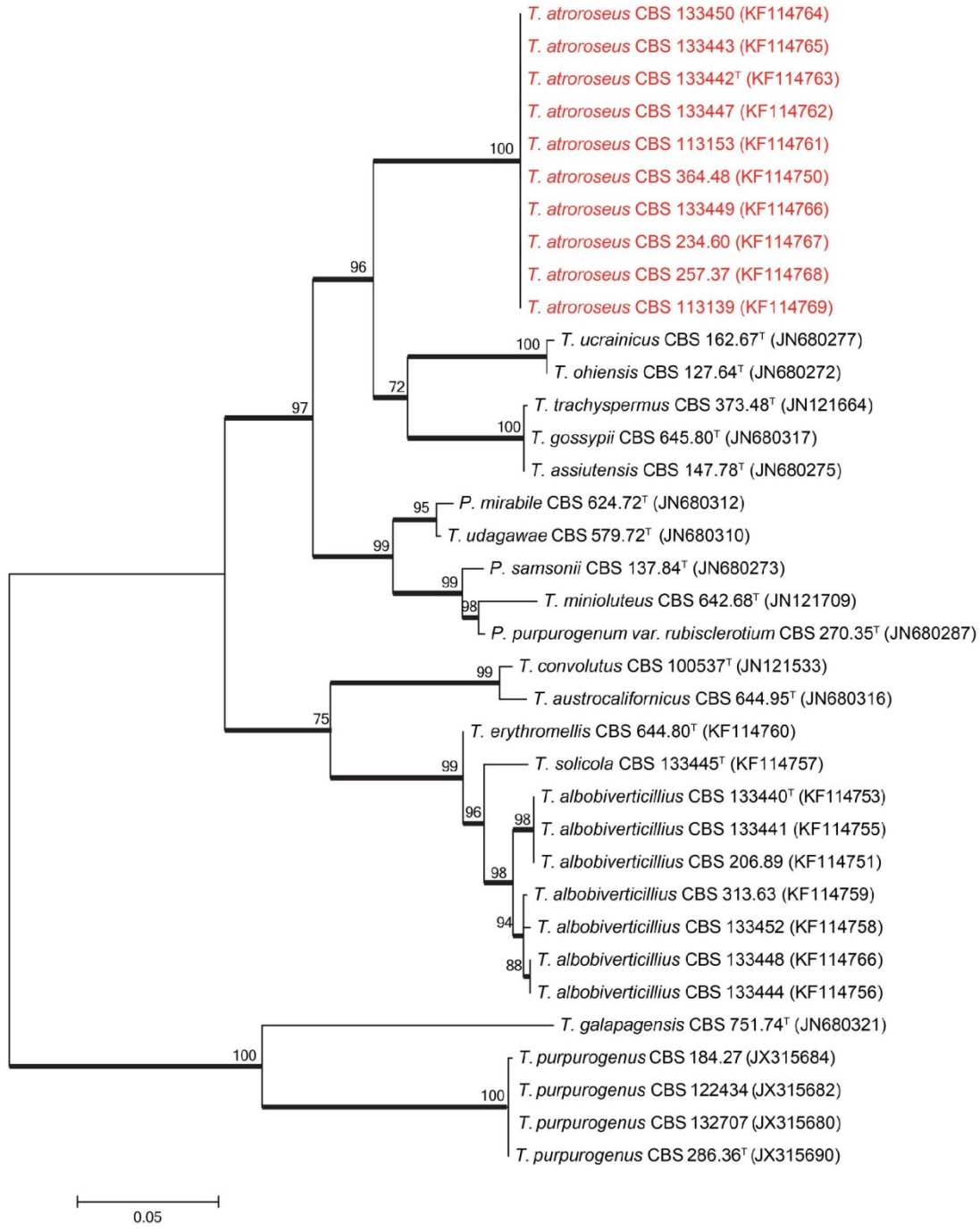


Figure 3. Maximum likelihood tree comparing the *RPB1* gene region of *Talaromyces* species closely related to *T. atroroseus*. *Talaromyces galapagensis* and *T. purpurogenus* were used as outgroup. Support in nodes is indicated above thick branches and is represented by bootstrap values higher than 70%. GenBank accession numbers are given between brackets, (T = ex-type). Red coloured names indicate *T. atroroseus* strains.

doi: 10.1371/journal.pone.0084102.g003

Table 2. Reported extrolite production by strains verified as *Talaromyces atroroseus* during this study.

Extrolite	Reported producer	Culture collection numbers	Reference
Glaucanic acid, Glauconic acid	*Penicillium "R. B.", P. purpurogenum*	**R.B.** = IMI 090178 = NRRL 1214 = CBS 113154 = IBT 3645 = IBT 4428	[23–30]
N-glutarylmonascorubramine, N-glutarylrubropunctamine	*P. purpurogenum*	**IBT 11181** = CBS 238.95 = CBS 123796	[9]
N-glutarylmonascorubramine	*P. purpurogenum*	R.B. = IMI 090178 = NRRL 1214 = CBS 113154 = **IBT 3645** = IBT 4428	[9]
Monascorubramine, PP-R	*P. purpurogenum*	**IBT 11180** = CBS 364.48 = ATCC 9777 = IMI 040037 = NRRL 1061 = QM 6760 = IBT 4458	[9]
PP-V, PP-R, PP-O, PP-Y	P. sp.	TA85S-28-H2 = **AZ** = IAM 15392 = JCM 23216 = IBT 32650	[43–47]
Purpuride	*P. purpurogenum*	**CBS 257.37**	[31]
Purpurogenone, Deoxypurpurogenone	*P. purpurogenum*	**CBS 257.37**	[32–35]
ZG-1494α	*P. rubrum*	IBT 11181 = **CBS 238.95** = CBS 123796	[36]

Strain numbers in bold are the strain numbers used in the references.

doi: 10.1371/journal.pone.0084102.t002

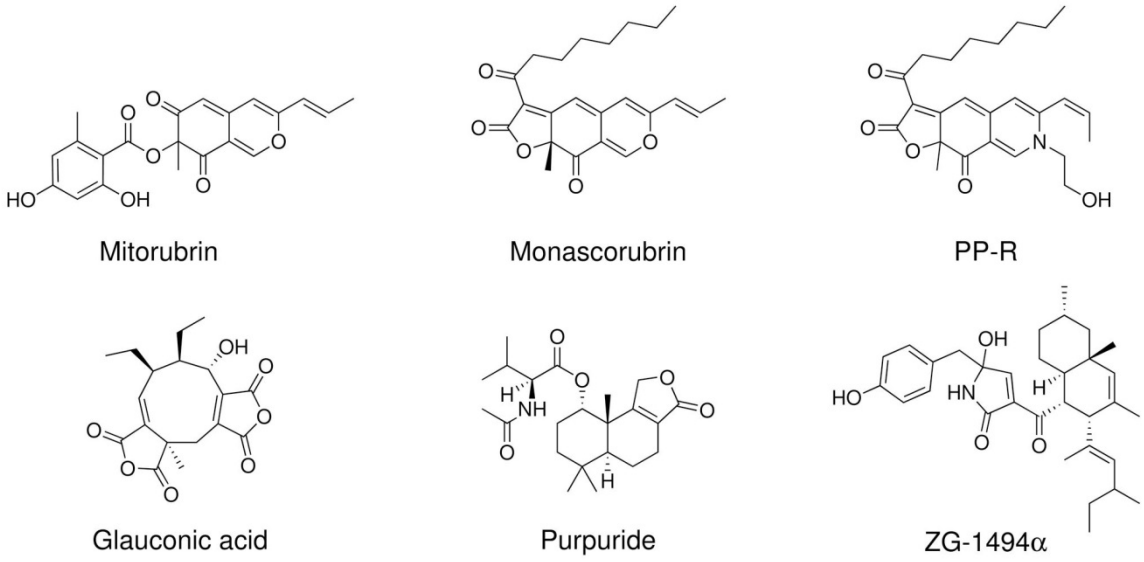


Figure 4. Structures of some of the most characteristic compounds produced by *Talaromyces atroroseus*. All six compounds were detected in this study.

doi: 10.1371/journal.pone.0084102.g004

Many *Talaromyces* species produce striking diffusing red pigments, especially *T. purpurogenus*, *T. atroroseus, T. albobiverticillius, T. minioluteus*, and *T. marneffei*. These red pigments are typically composed of the azaphilone pigments (Figure 5) monascorubrin, rubropunctatin, threonine derivative of rubropunctatin, monascorubramine, PP-R (= 7–(2–hydroxyethyl)-monascorubramine), rubropunctamine, N-glutarylrubropunctamine, and PP–V [8,9,43,44,61]. The same family of azaphilones are also known from red rice, where different species of *Monascus* have grown [1,2]. These red pigments are of interest for the industry as they are stable and non-toxic and can be used as food colorants [62]. The azaphilone pigments can react with amino acids, hence their name, and give intense dark red colours. In addition some of these species produce yellow azaphilone pigments, such as monascin, ankaflavin, monascusone A and B, xanthomonascin A, and another series of yellow mitrorubrin azaphilones: mitorubrin, mitorubrinol, mitorubrinol acetate, mitorubrinic acid, and many other related compounds [5]. Many of these pigments have been reported from or found in *T. atroroseus* in this study (Table 2 and Table 4). The potential for pigment production has in this study only been investigated in small scale on solid media; however, *T. atroroseus* also produce pigments in liquid cultures under the right conditions [8,46]. The

Table 3. Reported extrolite production from strains potentially belonging to *Talaromyces atroroseus*, but not examined during this study.

Extrolite	Reported producer	Strain identifier / Culture collection number	Reference
2,6,7-trihydroxy-3-methyl-naphthalene-1,4-dione	*Penicillium purpurogenum*	JS03-21*	[37]
BE-25327	*P. purpurogenum*	F25327 = FERM P-12345	[38]
Dhilirolide A, B, C, D	*P. purpurogenum*	IMI 357108	[39]
Glauconic acid	*P. glaucum*	-*	[40]
Gluconic acid	*P. purpurogenum* var. *rubrisclerotium* (= *T. pinophilus*)	No. 2670 = NRRL 1064 = CBS 270.35 = ATCC 4713 = ATCC 52224 = NRRL 1142 = IBT 4302	[41]
(-)-Mitorubrin,	*P. purpurogenum*	JS03-21*	[37]
Monascus red pigment	P. sp.	HKUCC 8070	[42]
Orsellinic acid	*P. purpurogenum*	JS03-21*	[37]
Purpactin A, B, C	*P. purpurogenum*	FO-608 = FERM P-10776	[48,49]
Purpurester A, B	*P. purpurogenum*	JS03-21*	[37]
Purpurquinones A, B, C	*P. purpurogenum*	JS03-21*	[37]
Red W59 ($C_{30}H_{34}O_9N_3$)	*P. purpurogenum*	-*	[50]
Red pigment	*P. purpurogenum*	GH2*	[51–53]
Red pigment	*P. purpurogenum*	SX01*	[54]
Red pigments	*P. purpurogenum*	DPUA 1275	[56,57]
Red pigments	*P. purpurogenum*	-*	[58,59]
Red pigments	P. sp.	-*	[60]
SL 3238 ($C_{27}H_{41}NO_7$)	*P. purpurogenum*	NRRL 3364	[55]
TAN-931	*P. purpurogenum*	JS03-21*	[37]

Based on the reported morphology and extrolites the strains in the table are by the authors' judgement belonging to *Talaromyces atroroseus* or a closely related species.

* Strain not deposited in any accessible culture collection

doi: 10.1371/journal.pone.0084102.t003

potential for up scaling the production of red pigments needs to be investigated thoroughly.

Even though sequence variations were observed for *Talaromyces albobiverticillius* strains, morphologically they were similar. Two strains used for the original description of *T. albobiverticillius* were received from Dr. Sung-Yuan Hsieh [63]. These included the type strain CBS 133440^T and CBS 133441. These strains were isolated from soil in Taiwan and produce white conidial masses and intense soluble red pigment on various media (Figure 6). However, other freshly isolated *T. albobiverticillius* strains produce densely sporulating colonies and do not show any stability for red pigment production. Some of the *Talaromyces albobiverticillius* strains did not produce any soluble pigment such as CBS 133444 and CBS 133448. Strains that did produce red pigments include CBS 113168, and CBS 133452. On MEA only the degraded or mutated strains of *T. albobiverticillius*, such as CBS 133440^T and CBS 313.63 produced red pigments. Micromorphologically all *T. albobiverticillius* strains produce long stipes (up to 380 µm) (Figure 5). Two strains of *T. albobiverticillius* (CBS 133440^T and CBS 133441) have globose to subglobose, smooth conidia; however, the remaining strains produce ellipsoid to fusiform smooth conidia (Figure 5).

Even though two clades were observed in the phylogenies there are no concordance between observed clades and morphological characters as discussed above. As such, they are considered here as representing one species. Raper and Thom [21] mentioned a number of colour mutations they observed in strains of *P. citrinum* and *P. chrysogenum*. They stated that colour mutations are encountered as the most common and conspicuous types of mutations, especially considering mature conidia. Mutations can often be observed when a strain loses its green pigment in its conidia, resulting in a white or tanned colour. Colour mutants are regularly encountered among the strains which were exposed to artificial stimulations such as ultra-violet, X-ray radiations and neutron bombardment [21].

Talaromyces atroroseus is considered as the optimal producer of industrially important yellow and red soluble pigments. Another option as a suitable producer of red soluble azaphilone pigments is *T. albobiverticillius*. However *T. albobiverticillius* produces soluble red pigment only in some strains. We speculate that the mitorubrins produced by *Talaromyces atroroseus* are of the (-)-form, as they have been shown to be that for the closely related *Talaromyces purpurogenus* (at that time identified as *Penicillium rubrum*) [64,65]. However, Natsume et al. [66] and Suzuki et al. [67] found both (+) and (-)-forms in the genus *Talaromyces*, while mitorubrins in *Hypoxylon* and other related genera are of the (+)-form [68–70]. Although *T. purpurogenus* is another good producer of diffusible red azaphilone pigments, this species also produce a series of mycotoxins, such as rubratoxin A and B and luteoskyrin in addition to extrolites that may be toxic if injected intraperitoneally (spiculisporic acid) [71] or in the veins of cats (rugulovasine A and B) [72,73]. *Talaromyces purpurogenus* can thus not be recommended for industrial production for red pigments.

Talaromyces atroroseus Yilmaz, Frisvad, Houbraken & Samson *sp. nov*. Figure 6.

Mycobank MB804901 [urn:lsid:mycobank.org: 804901]

Holotype: CBS 133442 in Centraalbureau voor Schimmelcultures is designated as the holotype of *Talaromyces atroroseus*. It was isolated from indoor house dust, Stellenbosch, South Africa by C. Visagie in 2010.

Cultures ex type: CBS 133442 = IBT 32470 = DTO 178A4 = KAS 3778

Etymology: Named after the dark rosy diffusing azaphilone pigment mixture produced.

Diagnosis: Dark green ellipsoidal rough-walled conidia and a dark red diffusing pigment, strains of the species produce the unique combination of secondary metabolites: glauconic acid, ZG–1494α, purpuride, red *Monascus* pigments, mitorubrins, and purpactins in fresh isolates.

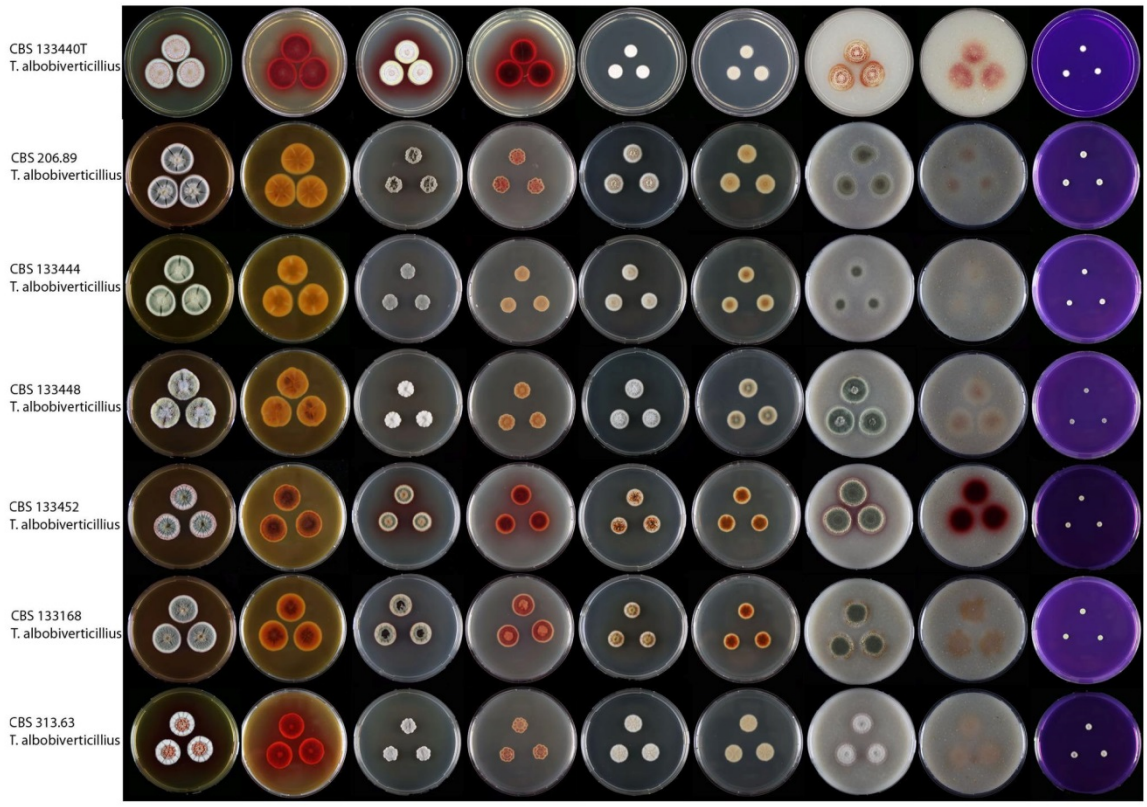


Figure 5. Strains of *Talaromyces albobiverticillius* on MEA, CYA, DG18, OA and CREA. Colony obverse and reverse is shown for the first four media and obverse for CREA.

doi: 10.1371/journal.pone.0084102.g005

CYA 25 °C 7d: Colonies are 30–40 mm in diameter, low, plane; margins narrow (1–2 mm), entire, low; mycelia white; texture velvety; sporulation dense, conidia *en masse* dark to dull green; exudate absent; soluble pigment red; reverse coloration dark cherry red.

MEA 25 °C 7d: Colonies 35–40 mm in diameter, low, plane, having a pinkish colour because of exudates diffusing into mycelia; margins narrow (1–2 mm), entire, low; mycelia white; texture velvety overlaying floccose; sporulation moderately dense, conidia *en masse* bluish green; exudate red droplets especially close to margin; soluble pigment absent, after prolonged incubation red pigments produced; reverse coloration dark red.

YES 25 °C 7d: Colonies are 33–45 mm in diameter, raised at centre, sulcate; margins wide (2–3 mm), entire, low; mycelia white; texture velvety; sporulation dense, conidia *en masse* dark to dull green; exudates small red droplets; soluble pigment red in some isolates; reverse coloration brownish red.

CYAS 25 °C 7d: Commonly no growth, some strains up to 5 mm in colony diameter.

CREA 25 °C 7d: Colonies 9–13 mm in diameter, weak acid production close to colony periphery, some strains acid absent; reverse dark red.

OA 25 °C 7d: Colonies 30–35 mm in diameter, low, plane; margins wide (2–3 mm), entire, low; white mycelia; texture velvety; sporulation dense; conidia *en masse* dull to dark green, almost appears blackish green; exudates absent; soluble pigment absent; reverse coloration commonly greenish yellow to green, red in some isolates.

DG18 25 °C 7d: Colonies 27–30 mm in diameter, low, plane; margins wide (2 mm), entire, low; mycelia white; texture velvety, floccose mycelia present at centre; sporulation dense, conidia *en masse* greyish green, at margins bluish green; exudates absent; soluble pigment absent; reverse colour is beige.

Conidiophores mostly biverticillate, subterminal branches produced, have a greenish to brownish pigmentation; Stipes smooth walled, 90–150 × 2.5–3 µm; Branches 2–3 when present, 15–50 × 2–3 µm; Metulae in verticils of 3 to 5 per stipe, 8–15 × 3.0–4.0 µm; Phialides acerose, 3 to 6 per metula,

Table 4. Extrolites of *Talaromyces atroroseus* and *T. albobiverticillius* as examined by HPLC-DAD and/or UHPLC- HRMS and comparison to standards on the media CYA and YES.

Species	Culture collection number	Extrolites* found
T. atroroseus	CBS 133450[a]	glauconic acid[b], monascorubrin[b], PP-R[b], purpuride[b], purpuroquinone A[b], ZG-1494α[b]
	CBS 113154[a]	glauconic acid, N-glutarylmonascorubramine[b], monascorubrin[b], PP-O[b], PP-R[b], purpuride[b], purpuroquinone A[b], ZG-1494α[b]
	CBS 123796[a]	FK17-P2b2[b], glauconic acid, N-glutarylmonascorubramine[b], mitorubrin, mitorubrinol, monascorubrin[b], PP-O[b], PP-R[b], purpuride[b], purpuroquinone A[b], purpurogenone, ZG-1494α[b]
	CBS 257.37	monascorubramine, purpuride, several Monascus-red pigments
	CBS 234.60	glauconic acid, monascorubramine, purpuride, ZG-1494α
	CBS 391.96	glauconic acid, monascorubramine, purpuride, ZG-1494α
	CBS 364.48	glauconic acid, monascorubramine, PP-R, purpuride, rubropunctatin, ZG-1494α
	CBS 133447	Glauconic acid, purpuride
	CBS 133442	Glauconic acid, monascorubramin, purpuride, rubropunctatin
	CBS 113153	glauconic acid, mitorubrin, monascorubramine, monascorubrin, purpuride
	CBS 113139	monascin, monascorubramine
	IBT 3933	glauconic acid, mitorubrin, monascorubramin, a purpactin
	IBT 20955	glauconic acid, monascorubramine, monascorubrin, purpuride, ZG-1494α
	IBT 23082	PP-R (only tested for *Monascus* pigments)
	CBS 133443	glauconic acid, monascorubramine, purpuride
	CBS 133449	glauconic acid, monascorubrin, purpuride
	JCM 23216	Glauconic acid, monascorubramine, purpuride
T. albobiverticillius	CBS 113168	mitorubrin, mitorubrinic acid, monascorubramine, PP-R, rubropunctatin, vermicellin
	CBS 313.63	mitorubrin, monascorubramin, monascorubrin, rubropunctatin
	IBT 4466	mitorubrinic acid, monascorubramine, a purpactin
	CBS 113167	mitorubrin, mitorubrinic acid, monascorubrin, a purpactin
	CBS 133444	mitorubrin, mitorubrinic acid, mitorubrinol
	CBS 133452	mitorubrin, mitorubrinic acid, monascorubramine, rubropunctatin
	CBS 133441	mitorubrin, mitorubrinic acid, monascin, monascorubramin, rubropunctatin, vermicellin

Table 4 (continued).

[a] Strains examined by both HPLC-DAD and UHPLC- -HRMS

[b] Extrolites identified by UHPLC- -HRMS

* The extrolites only identified by HPLC-DAD might in some cases not be the actual metabolite but a derivative with the same chromophore and retention on the column

doi: 10.1371/journal.pone.0084102.t004

9.5–12.5 × 2.5–3 µm; Conidia rough walled, ellipsoidal, 2–3.5 × 1.5–2.5 µm.

Talaromyces albobiverticillius (H.–M. Hsieh, Y.–M. Ju & S.–Y. Hsieh) Samson, Yilmaz, Frisvad & Seifert, Studies in Mycology 70: 174, 2011. MycoBank MB560683 (Figure 7)

Type. BCRC 34774

CYA 25 °C 7d: Colonies 15–20 mm in diameter, low, crateriform, in some isolates sulcate; margins narrow (1–2 mm), entire, low; mycelia white and yellow; texture floccose to velvety; sporulation sparse, in some isolates moderately dense; conidia *en masse* when sparse white, otherwise greyish green; exudates red small droplets; soluble pigmentation red; reverse coloration dark cherry red.

MEA 25 °C 7d: Colonies 24–28 mm in diameter, low, crateriform, in some isolates sulcate; margins wide (2–3 mm), entire, low; mycelia white and yellow; texture velvety with overlaying floccose in the centre; sporulation sparse, in some isolates moderately dense; conidia *en masse* when sparse white, otherwise greyish green; exudates clear and red droplets; soluble red pigment absent; reverse coloration dark red.

YES 25 °C 7d: Colonies 23–25 mm in diameter, raised at centre, sulcate; margins wide (2–3 mm), entire, low; mycelia white and yellow; texture velvety; sporulation sparse, in some isolates moderately dense; conidia *en masse* when sparse white, otherwise greyish green; exudates small orange to red droplets; soluble pigment red in some strains; reverse coloration red to pale brown.

CYAS 25 °C 7d: No growth.

CREA 25 °C 7d: Colonies 4–8 mm in diameter, no acid produced.

OA 25 °C 7d: Colonies 25–28 mm in diameter, low, plane; margins wide (3–4 mm), entire, low; mycelia white; texture velvety; sporulation sparse to moderately dense; conidia *en masse* when sparse white, otherwise greyish green; exudates absent; soluble pigment absent; reverse coloration red in the centre and the rest greenish yellow to green.

DG18 25 °C 7d: Colonies 15–35 mm in diameter, low, plane; margins narrow (1–2 mm), entire, low; mycelia white; texture velvety; sporulation sparse; sparse to moderately dense; conidia *en masse* when sparse white, otherwise greyish green; exudates clear to red droplets; soluble pigment red in some isolates absent; reverse coloration brownish red, in some isolates beige.

Conidiophores strictly biverticillate, subterminal branches absent; stipes smooth walled, 200–380 × 2.5–3.5 µm; metulae in verticals of 3–6, 8–12 × 1.5–4.5 µm; phialides acerose, 3–7 per metula, 8–13.5 × 2–3 µm; conidia smooth to finely

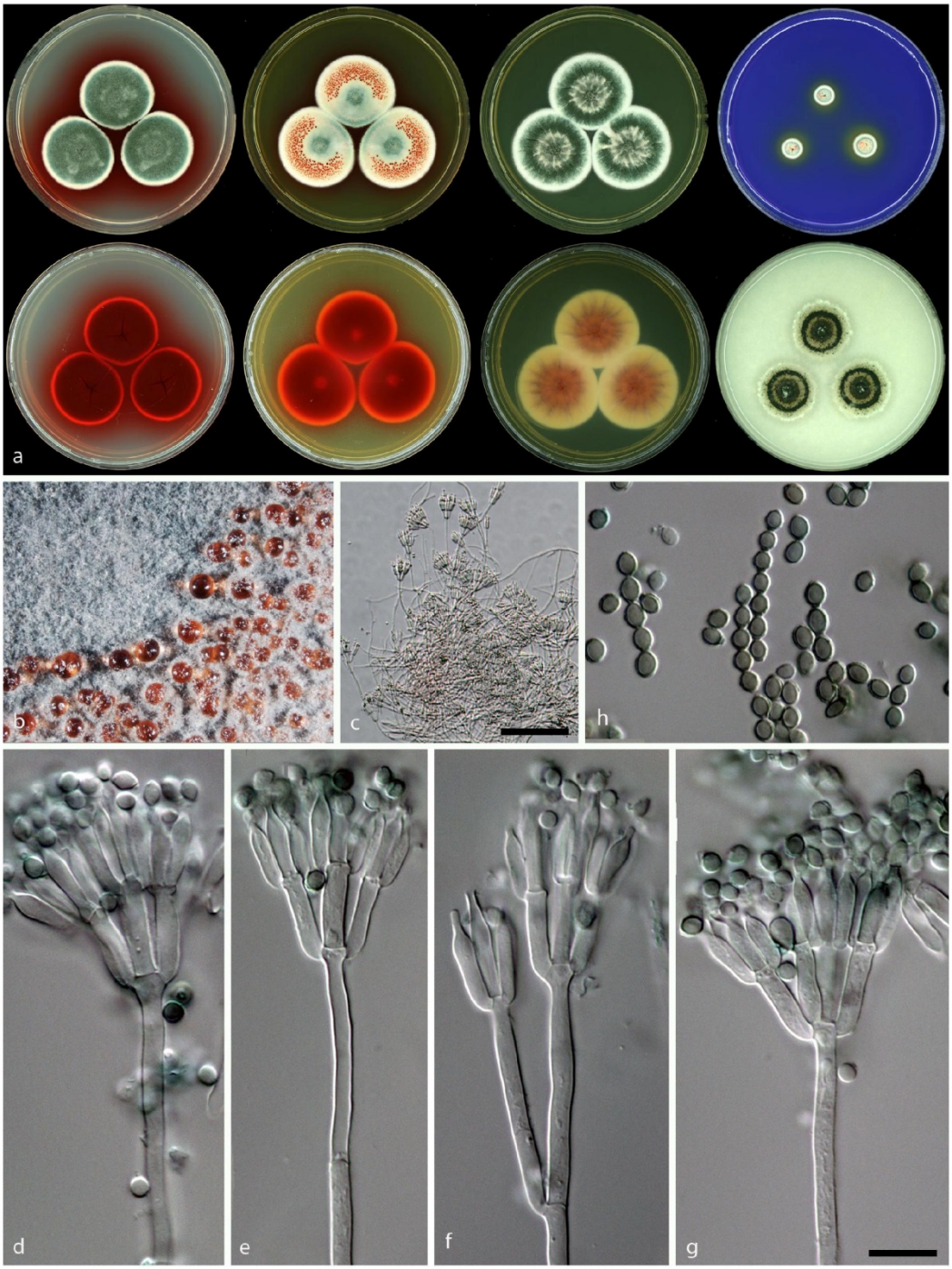


Figure 6. Morphological features of *Talaromyces atroroseus* sp. nov. CBS 133442. a: Colonies incubated on CYA, CYA reverse, MEA, MEA reverse, YES, YES reverse, CREA and OA from left to right b: Colony texture on MEA, c–g: Conidiophores produced on MEA; h: Conidia. (– Scale Bar in c = 50 µm, in g = 10 µm and applies to d–h).

doi: 10.1371/journal.pone.0084102.g006

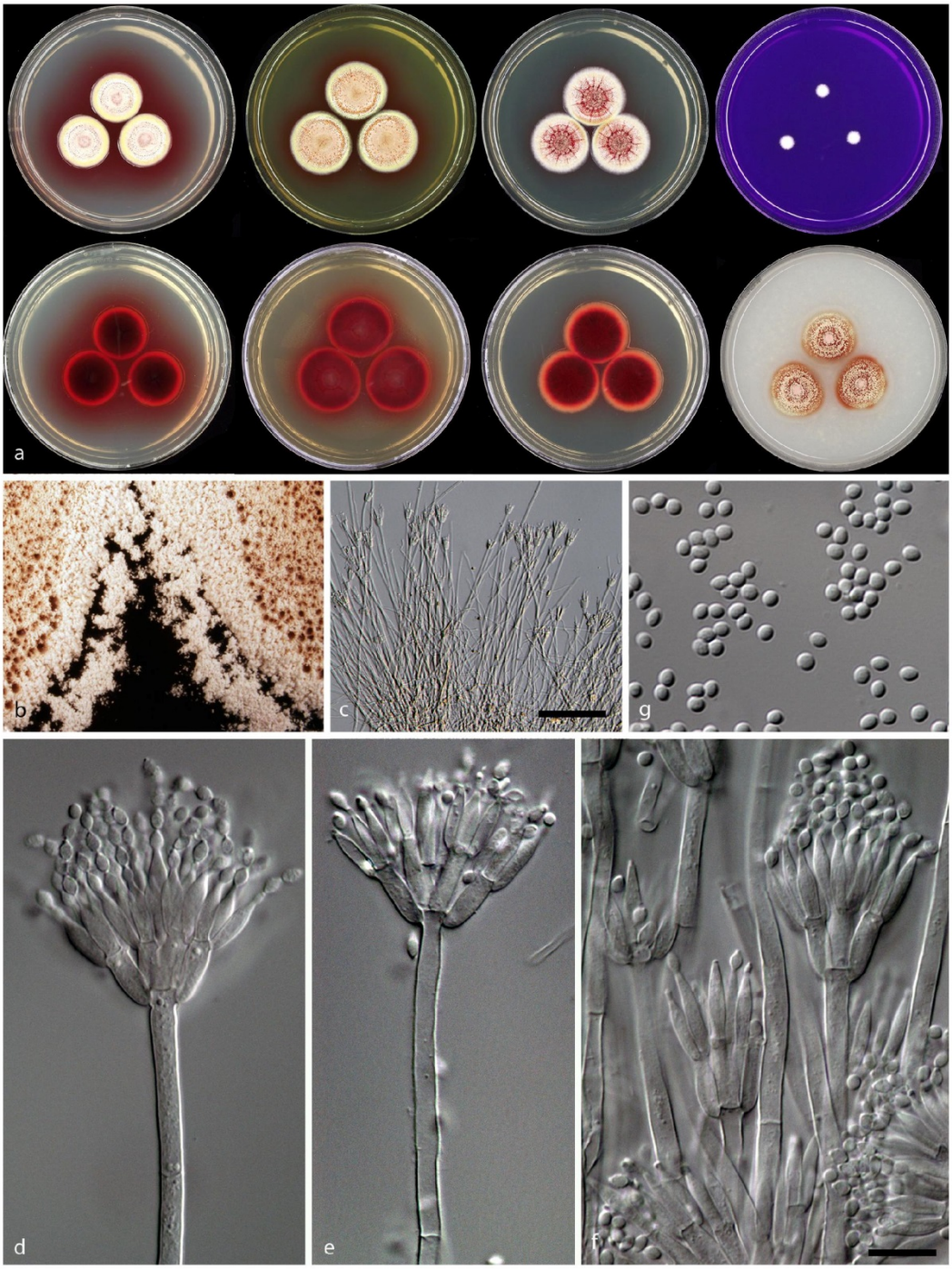


Figure 7. Morphological features of *Talaromyces albobiverticillius* CBS 133440. : Colonies incubated on CYA, CYA reverse, MEA, MEA reverse, YES, YES reverse, CREA and OA from left to right b: Colony texture on MEA, c–f: Conidiophores produced on MEA; g: Conidia. (– Scale Bar in c = 50 µm, in f = 10 µm and applies to d–g).

doi: 10.1371/journal.pone.0084102.g007

roughened, spheroid to subglobose, in some isolates fusiform, 2–3.5 (4) × 1.5–2.5 µm.

Conclusion

Talaromyces atroroseus is a new species that produce large amounts of red pigments that can be potentially used for colouring foods, as it does not produce any known mycotoxins. Certain strains of *T. albobiverticillius* may also be used for these purposes.

Supporting Information

Table S1. Table S1 contains the extrolites searched for by ultra high performance-liquid chromatography-diode array detection-high resolution mass spectrometric detection (UHPLC-DAD-HRMS) the fungal extracts analysed. The table also includes data on the available standards used in the study.

(DOCX)

Acknowledgements

We thank Agilent Technologies for the donation of LC-QTOF and LC-QQQ equipment to DTU Systems Biology via a Thought Leader Award. We also thank Cobus Visagie and Keith A. Seifert for a strain of *Talaromyces atroroseus* (KAS 3778) that served as type of the new species an was isolated from an indoor mold project funded by the Alfred P. Sloan Foundation. Ellen Kirstine Lyhne is thanked for technical assistance.

Author Contributions

Conceived and designed the experiments: JCF NY JH RAS. Performed the experiments: JCF NY KBR JH. Analyzed the data: JCF NY UT KBR JH RAS. Contributed reagents/ materials/analysis tools: JCF NY KBR. Wrote the manuscript: JCF NY. Read and improved the draft paper: JCF NY UT KBR JH RAS.

References

1. Feng Y, Shao Y, Chen F (2012) *Monascus* pigments. Appl Microbiol Biotechnol 96: 1421-1440. doi:10.1007/s00253-012-4504-3. PubMed: 23104643.
2. Patakova P (2013) *Monascus* secondary metabolites: production and biological activity. J Ind Microbiol Biotechnol 40: 169-181. doi:10.1007/s10295-012-1216-8. PubMed: 23179468.
3. Samson RA, Stolk AC, Frisvad JC (1989) Two new synnematous species of *Penicillium*. Stud Mycol 31: 133-143.
4. van Reenen-Hoekstra ES, Frisvad JC, Samson R.A, Stolk AC (1990). The *Penicillium funiculosum* complex - well defined species and problematic taxa. In: RA , Pitt JI (eds.): Modern concepts in *Penicillium* and *Aspergillus* classification. New York: Plenum Press. pp. 173-191
5. Frisvad JC, Filtenborg O, Samson RA, Stolk AC (1990) Chemotaxonomy of the genus *Talaromyces*. Antonie Van Leeuwenhoek 57: 179-189. doi:10.1007/BF00403953. PubMed: 2181929.
6. Lin YL, Wang TH, Lee MH, Su NW (2008) Biologically active components and nutraceuticals in the *Monascus*–fermented rice: a review. Appl Microbiol Biotechnol 77: 965-973. doi:10.1007/s00253-007-1256-6. PubMed: 18038131.
7. Liu BH, Wu TS, Su MC, Chung CP, Yu FY (2005) Evaluation of citrinin occurrence and cytotoxicity in *Monascus* fermentation products. J Agric Food Chem 53: 170-175. doi:10.1021/jf048878n. PubMed: 15631525.
8. Mapari SAS, Hansen ME, Meyer AS, Thrane U (2008) Computerized screening for novel producers of *Monascus*–like pigments in *Penicillium* species. J Agric Food Chem 56: 9981–9989. doi:10.1021/jf801817q. PubMed: 18841978.
9. Mapari SAS, Meyer AS, Thrane U, Frisvad JC (2009) Identification of potentially safe promising fungal cell factories for the production of polyketide natural food colorants using chemotaxonomic rationale. Microb Cell Fact 8: 24. doi:10.1186/1475-2859-8-24. PubMed: 19397825.
10. Yilmaz N, Houbraken J, Hoekstra ES, Frisvad JC, Visagie CM et al. (2012) Delimitation and characterisation of *Talaromyces purpurogenus* and related species. Persoonia 29: 39-54. doi: 10.3767/003158512X659500. PubMed: 23606764.
11. Samson RA, Houbraken J, Thrane U, Frisvad JC, Andersen B (2010) Food and indoor fungi. CBS laboratory manual series 2. Utrecht: CBS KNAW Fungal Biodiversity Centre. p. 390.
12. Houbraken J, Due M, Varga J, Meijer M, Frisvad JC et al. (2007) Polyphasic taxonomy of *Aspergillus* section *Usti*. Stud Mycol 59: 107-128. doi:10.3114/sim.2007.59.12. PubMed: 18490949.
13. Houbraken J, López Quintero CA, Frisvad JC, Boekhout T et al. (2011) Five new *Penicillium* species, *P. araracuarense*, *P. elleniae*, *P.* penarojaense, P. vanderhammenii and P. wotroi, from Colombian leaf litter. Int J Syst Evol Microbiol 61: 1462-1475.
14. Houbraken J, Samson RA (2011) Phylogeny of *Penicillium* and the segregation of *Trichocomaceae* into three families. Stud Mycol 70: 1–51. doi:10.3114/sim.2011.70.01. PubMed: 22308045.
15. Samson RA, Yilmaz N, Houbraken J, Spierenburg H, Seifert KA et al. (2011) Phylogeny and nomenclature of the genus *Talaromyces* and taxa accommodated in *Penicillium* subgenus *Biverticillium*. Stud Mycol 70: 159-183. doi:10.3114/sim.2011.70.04. PubMed: 22308048.
16. Tamura K, Peterson D, Peterson N, Stecher G, Nei M et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731-2739. doi:10.1093/molbev/msr121. PubMed: 21546353.
17. Frisvad JC, Thrane U (1987) Standardized High-Performance Liquid Chromatography of 182 mycotoxins and other fungal metabolites based on alkylphenone indices and UV-VIS spectra (diode-array detection). J Chromatogr 404: 195-214. doi:10.1016/S0021-9673(01)86850-3. PubMed: 3680432.
18. Nielsen KF, Månsson M, Rank C, Frisvad JC, Larsen TO (2011) Dereplication of microbial natural products by LC–DAD–TOFMS. J Nat Prod 74: 2338-2348. doi:10.1021/np200254t. PubMed: 22026385.
19. Houbraken J, Spierenburg H, Frisvad JC (2012) *Rasamsonia*, a new genus comprising thermotolerant and thermophilic *Talaromyces* and *Geosmithia* species. Antonie Van Leeuwenhoek 101: 403-421. doi: 10.1007/s10482-011-9647-1. PubMed: 21965082.
20. Stoll O (1903–1904) Beiträge zur Morphologischen und Biologischen Charakteristik von *Penicillium*–Arten. Dissertation Würzberg. 56 p.
21. Raper KB, Thom C (1949) Manual of Penicillia. Baltimore, MD, USA: Williams & Wilkins. 875 pp.
22. Pitt JI (1980) The genus *Penicillium* and its teleomorphic states *Eupenicillium* and *Talaromyces*. New York: Academic Press. p. 634.
23. Yuill JL (1934) The acids produced from sugar by a *Penicillium* parasitic upon *Aspergillus niger*. Biochem J 28: 222-227. PubMed: 16745356.
24. Takashima M, Kitajama A, Otuka K (1955) Studies on the pigment of a white crystal "glauconic acid" and red pigments by *Penicillium purpurogenum*. J Agric Chem Soc JAPAN 29: 25-29.
25. Baldwin JE, Barton DHR, Bloomer JL, Jackman LM, Rodriguez-Hahn L et al. (1962) The constitutions of glauconic, glaucanic and byssochlamic acids. Experientia 18: 345-352. doi:10.1007/BF02172243. PubMed: 13864341.
26. Barton DHR, Jackman LM, Rodriguez-Hahn L, Sutherland JK (1965) The nonadrides. Part. II. The constitutions of glauconic and glaucanic acids. J Chem Soc 1965: 1772-1778
27. Barton DHR, Godinho LDS, Sutherland JK (1965) The nonadrides. Part. III. The absolute configuration of glauconic and glaucanic acids. J Chem Soc 1965: 1779-1786

28. Barton DHR, Sutherland JLK (1965) The nonadrides. Part I. Introduction and general survey. J Chem Soc 1965: 1769-1772
29. Huff RK, Moppett CE, Sutherland JK (1968) A novel synthesis of a nine-membered ring. J Chem Soc Chem Commun 1968: 1192-1193
30. Huff RK, Moppett CE, Sutherland JK (1972) Nonadrides 6. Dimerization of C-9 unit *in vivo* and *in vitro*. J Chem Soc 1972: 2584-2590
31. King TJ, Roberts JC, Thompson DJ (1973) Studies in mycological chemistry. Part XXX and last. Isolation and structure of purpuride, a metabolite of Penicillium purpurogenum Stoll. J Chem Soc Perkin I 1973: 78-80.
32. Roberts JC, Warren CWH (1955) Studies in mycological chemistry. Part IV. Purpurogenone, a metabolic product of *Penicillium purpurogenum* Stoll. J Chem Soc 1955: 2992-2998
33. King TJ, Roberts JC, Thompson DJ (1970) The structure of purpurogenone, a metabolite of *Penicillium purpurogenum* Stoll, an X-ray study. J Chem Soc D: 1499.
34. Roberts JC, Thompson DJ (1971) Studies in mycological chemistry. Part XXVII. Reinvestigation of the structure of purpurogenone, a metabolite of *Penicillium purpurogenum* Stoll. J Chem Soc 1971: 3488-3492
35. Roberts JC, Thompson DJ (1971) Studies in mycological chemistry. Part XXVIII. Isolation and structure of deoxypurpurogenone, a minor pigment of *Penicillium purpurogenum* Stoll. J Chem Soc 1971: 3493-3495
36. West RR, van Ness J, Varming A-M, Rassing B, Biggs S et al. ZG (1494α) a novel platelet-activating acyltransferase inhibitor from *Penicillium rubrum*, isolation, structure elucidation and biological activity. J Antibiot 49: 967-973.
37. Wang H, Wang Y, Wang W, Fu P, Liu P et al. (2011) Anti-influenza virus polyketides from the acid-tolerant fungus *Penicillium purpurogenum* JS03-21. J Nat Prod 74: 2014-2018. doi:10.1021/np2004769. PubMed: 21879714.
38. Kondo H, Kurama M, Nakajima S, Osada K, Okura A et al. (1993) New substance BE-25327 which is an estradiol agonist – useful for treating oestrogen-deficient gynaecological diseases, osteoporous, prostatic cancer and prostatomegaly. Jpn. Kokai Tokkyo Koho, JP05032579 A2, 9 Feb
39. Silva ED, Williams DE, Jayanetti DR, Centko RM, Patrick BO et al. (2011) Dhilirolides A-D, meroterpenoids produced in culture by the fruit-infecting fungus *Penicillium purpurogenum* collected in Sri Lanka. Org Lett 13: 1174-1177. doi:10.1021/ol200031t. PubMed: 21348535.
40. Wijkman N (1931) On some new substances made through mould fungus. Justus. Liebigs Ann Chem 485: 61-73. doi:10.1002/jlac.19314850106.
41. Herrick HT, Mayo E (1928) The production of gluconic acid by the *Penicillium luteum-purpurogenum* group. II. Some optimal conditions for acid formation. J Biol Chem 77: 185-195.
42. Jiang Y, Li HB, Chen F, Hyde KD (2005) Production potential of water-soluble *Monascus* red pigment by a newly isolated *Penicillium* sp. J Agric Technol 1: 113-126.
43. Ogihara J, Kato J, Oishi K, Fujimoto Y, Eguchi T (2000) Production and structural analysis of PP-V, a homologue of monascorubramine, produced by a new isolate of *Penicillium* sp. J Biosci Bioeng 90: 549-554. doi:10.1016/S1389-1723(01)80039-6. PubMed: 16232908.
44. Ogihara J, Kato J, Oishi K, Fujimoto Y (2001) PP-R, 7-(2-hydroxyethyl)-monascorubramine, a red pigment produced in the mycelia of *Penicillium* sp. AZ. J Biosci Bioeng 91: 44-47. doi:10.1263/jbb.91.44. PubMed: 16232944.
45. Ogihara J, Kato J, Oishi K (2002) Effect of ammonium nitrate on the production of PP-V and monascorubrin homologues by *Penicillium* sp. AZ. J Biosci Bioeng 93: 54-59. doi:10.1263/jbb.93.54. PubMed: 16233165.
46. Arai T, Umemura S, Ota T, Ogihara J, Kato J et al. (2012) Effects of inorganic nitrogen sources on the production of PP-V [(10Z)-12-carboxyl-monascorubramine] and the expression of the nitrate assimilation gene cluster by *Penicillium* sp. AZ. Biosci Biotechnol Biochem 76: 120-124. doi:10.1271/bbb.110589. PubMed: 22232256.
47. Arai T, Koganei K, Umemura S, Kojima R, Kato J et al. (2013) Importance of the ammonia assimilation by *Penicillium purpurogenum* in amino derivative *Monascus* pigment, PP-V, production. AMB Express 3: 19. doi:10.1186/2191-0855-3-19. PubMed: 23537394.
48. Tomoda H, Nishida H, Masuma R, Cao J, Okuda S et al. (1991) Purpactins, new inhibitors of acyl-CoA:cholesterol acyltransferase produced by *Penicillium purpurogenum*. I. Production, isolation and physico-chemical and biological properties. J Antibiot (Tokyo) 44: 136-143. doi:10.7164/antibiotics.44.136. PubMed: 1750931.
49. Nishida H, Tomoda H, Cao J, Okuda S, Omura S (1991) Purpactins, new inhibitors of acyl-CoA:cholesterol acyltransferase produced by *Penicillium purpurogenum*. II. Structure elucidation of purpactins A, B and C. J Antibiot (Tokyo) 44: 144-151. doi:10.7164/antibiotics.44.144. PubMed: 2010354.
50. Pharm Amano (1974) Novel pigment Red W59 extracted from *Penicillium purpurogenum* culture medium with organic solvents, Jap. Pat. JP 49093587-A
51. Méndez-Zavala A, Contreras-Esquivel JC, Lara-Victoriano F, Rodriguez-Herrera R, Aguilar CN (2007) Fungal production of the red pigment using a xerophilic strain *Penicillium purpurogenum* GH-2. Rev Mex Ingen Quim 6: 267-273.
52. Méndez A, Pérez C, Montañéz JC, Martínez GAguilar CN (2011) Red pigment production by *Penicillium purpurogenum* GH2 is influenced by pH and temperature. J Zhejiang Univ-Sci B (Biomed & Biotechnol) 12: 961-968
53. Espinoza-Hernández TC, Rodríguez-Herrera R, Aguilar-González CN, Lara-Victoriano F, Reyes-Valdés MH et al. (2013) Characterization of three novel pigment-producing *Penicillium* strains isolated from the Mexican semi-desert. Afr J Biotechnol 12: 3405-3413.
54. Qui M, Xie R, Shi Y, Chen H, Wen Y et al. (2010) Isolation and identification of endophytic fungus SX01, a red pigment producer from *Ginkgo biloba* L. World J Microbiol Biotechnol 26: 993-998. doi: 10.1007/s11274-009-0261-6.
55. Bollinger P, Haerri E, Sigg H (1972) Antibiotic derived from *Penicillium* species. US Pat. 3,655,880
56. Santos-Ebinuma VC, Teixeira MFS, Pessoa A Jr. (2013) Submerged culture conditions for the production of alternative natural colorants by a new isolated *Penicillium purpurogenum* DPUA 1275. J Microbiol Biotechnol 23: 802-810. doi:10.4014/jmb.1211.11057. PubMed: 23676916.
57. Santos-Ebinuma VC, Roberto IC, Teixeira MFS, Pessoa A Jr. (2013) Improving of red colorants production by a new *Penicillium purpurogenum* strain in submerged culture and the effect of different parameters in their stability. Biotechnol Prog 29: 778-785. doi:10.1002/btpr.1720. PubMed: 23554384.
58. Velmurugan P; Kamala, Kannan S, Balachandar V, Lakshmanaperumalsamy P, Chae J-C et al. (2010) Natural pigment extraction from five filamentous fungi for industrial applications and dyeing of leather. Carb Polym 79: 262-268. doi:10.1016/j.carbpol.2009.07.058. Available online at: doi:10.1016/j.carbpol.2009.07.058
59. Velmurugan P, Lee YH, Venil CK, Lakshmanaperumalsamy P, Chae J-C et al. (2010) Effect of light on growth, intra cellular and extracellular pigment production by five pigment-producing filamentous fungi in synthetic medium. J Biosci Bioeng 109: 346-350. doi:10.1016/j.jbiosc.2009.10.003. PubMed: 20226375.
60. Gunasekaran S, Poorniammai R (2008) Optimization of fermentation conditions for red pigment production from *Penicillium* sp. under submerged cultivation. Afr J Biotechnol 7: 1894-1898.
61. Mapari SAS, Nielsen KF, Larsen TO, Frisvad JC, Meyer AS et al. (2005) Exploring fungal biodiversity for water–soluble pigments and potential natural food colorants. Current Opinion Biotechnol 16: 231-238. doi:10.1016/j.copbio.2005.03.004.
62. Mapari SAS, Thrane U, Meyer AS (2010) Fungal polyketide azaphilone pigments as future food colorants? Trends Biotechnol 28: 300-307. doi: 10.1016/j.tibtech.2010.03.004. PubMed: 20452692.
63. Hsieh HM, Ju YM, Hsieh SY (2010) *Penicillium albobiverticillium* sp. nov., a new species producing white conidial masses from biverticillate penicillia. Fung Sci 25: 25-31.
64. Büchi G, White JD, Wogan GN (1965) The structures of mitorubrin and mitorubrinol. J American Chem Soc 87: 3484-3489. doi:10.1021/ja01093a036.
65. Marsini MA, Gowin KM, Pettus TRR (2006) Total synthesis of (+)-mitorubrinic acid. Org Lett 8: 3481-3483. doi:10.1021/ol0610993. PubMed: 16869640.
66. Natsume M, Takahashi Y, Marumo S (1985) (-)-Mitorubrinic acid, a morphogenic substance inducing chlamydospore-like cells, and its related new metabolite, (+)-mitorubrinic acid B, isolated from *Penicillium funiculosum*. Agric Biol Chem 49: 2517-2519. doi:10.1271/bbb1961.49.2517.
67. Suzuki S, Hosoe T, Nozawa K, Yaguchi T, Udagawa Si et al. (1999) Mitorubrin derivatives on ascomata of some *Talaromyces* species of Ascomycetous fungi. J Nat Prod 62: 1328-1329. doi:10.1021/np990146f. PubMed: 10514327.
68. Steglich W, Klaar M, Furtner W (1974) (+)-Mitorubrin derivatives from *Hypoxylon fragiforme*. Phytochem: 132874-132875.
69. Osmanova N, Schultze W, Ayoub N (2010) Azaphilones: a class of fungal metabolites with diverse biological activities. Phytochem Rev 9: 315-342. doi:10.1007/s11101-010-9171-3.
70. Gao J-M, Yang S-X, Qin J-C (2013) Azaphilones. Chemistry and Biology - Chem Rev 113: 4755-4811.

71. Fujimoto H, Jisai Y, Horie Y, Yamazaki M (1988) On isolation of spiculisporic acid, a toxic metabolite from *Talaromyces panasenkoi*. Proc Jpn Assoc Mycotoxicol 27. pp. 15-19.
72. Nagaoka A, Kukuchi K, Nagawa Y (1972) Pharmacological studies of new indole alkaloids, rugulovasine A and B hydrochloride. I. Effects of both alkaloids on cardiovascular and central nervous system, and smooth muscles. Drug Res 22: 137-142.
73. Nagaoka A, Kikuchi K (1972) Pharmacological studies of new indole alkaloids, rugulovasine A and B hydrochloride. II. Hypotensive mechanism of both alkaloids in the anesthetized cats. Drug Res 22: 143-146.

A.2 Supplementary material

Talaromyces atroroseus, a new species efficiently producing industrially relevant red pigments.

by

Jens C. Frisvad, Neriman, Yilmaz, Ulf Thrane, Kasper Bøwig Rasmussen, Jos Houbraken, Robert A. Samson

Table A.1. Library of compounds searched for by UHPLC-DAD-HRMS including available standards. The table contains the used for targeting search on accurate mass. The observed adducts and losses confirming the identity of a compound have been reported for standards as well as identified compound in this study. Furthermore it is stated where UV/VIS spectra have been used in the confirmation of compound identity. RT: Retention time relating to the UHPLC method described.

Compound	Formula	Monoisotopic Mass	Adducts/Losses observed, ESI+	RT	Confirmed by UV/VIS	Standard Available
Ankaflavin	C23H30O5	386.2093	$[M+H]^+$=387.2169, $[M+Na]^+$=409.1988	12.418		Yes
Deoxypurpurogenone	C29H20O10	528.1056	$[M+H]^+$=529.1134	11.299	Yes	No
Dhilirolide A	C25H28O9	472.1733				No
Dhilirolide B	C25H28O8	456.1784				No
Dhilirolide C	C25H28O8	456.1784				No
Dhilirolide D	C25H30O7	442.1992				No
FK17-P2b2	C13H16O4	236.1049	$[M+H]^+$=237.1130, $[M+Na]^+$=259.0952, $[M-H_2O+H]^+$= 219.1026	6.150		No
Glauconic acid	C18H20O7	348.1209	$[M+H]^+$=349.1284, $[M+Na]^+$=371.1105, $[M+NH_4]^+$=366.1551, $[M-H_2O+H]^+$= 331.1176	7.991 / 8.846		Yes
(+)-Mitorubrin	C21H18O7	382.1053	$[M+H]^+$=383.1127, $[M+Na]^+$=405.0951, $[M-H_2O+H]^+$= 365.1024	9.478	Yes	Yes
(+)-Mitorubrinic acid	C21H16O9	412.0794	$[M+H]^+$=413.0873, $[M+Na]^+$=435.0687, $[M-H_2O+H]^+$= 395.0763	7.198	Yes	Yes
(+)-Mitorubrinol	C21H18O8	398.1002	$[M+H]^+$=399.1079, $[M+Na]^+$=421.0894, $[M-H_2O+H]^+$= 281.0964	7.126	Yes	Yes
(+)-Mitorubrinol acetate	C23H20O9	440.1107	$[M+H]^+$=441.1187, $[M+Na]^+$=463.1003	8.846	Yes	Yes
Monaphilone A	C22H32O4	360.2301				No
Monaphilone B	C20H28O4	332.1988				No
Monaphilone C	C20H32O4	336.2301				No
Monascin	C21H26O5	358.1780	$[M+H]^+$=359.1851, $[M+Na]^+$=381.1666	11.438	Yes	Yes
Monascorubramine	C23H27NO4	381.1940	$[M+H]^+$=382.2015, $[M+Na]^+$=404.1836	10.307	Yes	Yes
Monascorubrin	C23H26O5	382.1780	$[M+H]^+$=383.1856, $[M+Na]^+$=405.1675	12.528	Yes	Yes
Monascusone A	C13H18O5	254.1154				No
Monascusone B	C17H18O5	302.1154				No

Monascuspurpurone	C21H30O5	362.2093				No
Monasfluore A	C21H24O5	356.1624				No
Monasfluore B	C23H28O5	384.1937				No
N-glutarylmonascorubramine	C28H33NO8	511.2206	$[M+H]^+$=512.2279, $[M+Na]^+$=534.2100	9.521	Yes	No
N-glutarylrubropunctamine	C26H29NO8	483.1893				No
PP-O	C23H24O7	412.1522	$[M+H]^+$=413.1598, $[M+Na]^+$=435.1415	10.761	Yes	No
PP-R	$C_{25}H_{31}NO_5$	425.2202	$[M+H]^+$=426.2279, $[M+Na]^+$=448.2096	10.050	Yes	No
PP-V	C23H25NO6	411.1682				
Purpactin A	C23H26O7	414.1679				No
Purpactin B	C23H26O7	414.1679				No
Purpactin C	C23H24O7	412.1522				No
Purpurester A	C13H16O5	252.0998				No
Purpurester B	C12H12O4	220.0736				No
Purpuride	C22H33NO5	391.2359	$[M+H]^+$=392.2429, $[M+NH_4]^+$= 409.2696, $[M+Na]^+$=414.2256	9.110	Yes	No
Purpurogenone	C29H20O11	544.1006	$[M+H]^+$=545.1079	11.106	Yes	Yes
Purpurquinone A	C21H20O9	416.1107	$[M+H]^+$=417.1186, $[M+Na]^+$=439.1004	7.383	Yes	No
Purpurquinone B	C21H20O8	400.1158				No
Purpurquinone C	C21H20O10	432.1056				No
Rubratoxin A	C26H32O11	520.1945				Yes
Rubratoxin B	C26H30O11	518.1788				Yes
Rubropunctamine	C21H23NO4	353.1627	$[M+H]^+$=354.1704 $[M+Na]^+$=376.1523	8.961	Yes	Yes
Rubropunctatin	C21H22O5	354.1467		11.489		Yes
Rugulovasine A	C16H16N2O2	268.1212				No
Rugulovasine B	C16H16N2O2	268.1212				Yes
Xanthomonasin A	C21H24O7	388.1522				No
Xanthomonasin B	C23H28O7	416.1835				No
ZG-1494a	C32H43NO4	505.3192	$[M+H]^+$=506.3267, $[M+Na]^+$=528.3093, $[M-H_2O+H]^+$= 488.3161	12.009		No

Appendix B

Supporting information to Chapter 4

Isolation and structural characterisation of a novel amino acid derivative *Monascus* pigment, N-glutamyl monascorubraminic acid, from *Talaromyces atroroseus* IBT11181

Thomas Isbrandt[1], Kasper Bøwig Rasmussen[1], Gerit Nymschefsky[1], Evgenia Tsanaktsidou[1], Ulf Thrane[1], Thomas Ostenfeld Larsen[1]

[1] Section for Eukaryotic Biotechnology, Department of Systems Biology, Technical University of Denmark, Søltofts Plads B221, DK-2800 Kgs. Lyngby, Denmark

This is a preliminary manuscript; further studies will be conducted before publication.

B.1 Abstract

In the present study a novel azaphilone type of compound is isolated from liquid fermentation of a *Talaromyces atroroseus* IBT11181 and structure elucidated. The compound is the glutamine derivative of PP-V and is provisionally named N-glutamyl monascorubraminic acid.

B.2 Experimental

Talaromyces atroroseus IBT11181, retrieved from the the IBT collection at Department of Systems biology, Technical University of Denmark, Lyngby, Denmark, was cultivated in 2 L Satorius Biostat B reactors with a working volume of 1.8 L. The bioreactors were equipped with two Rushton six-blade disc turbines for proper agitation. The pH was monitored with a calibrated pH electrode (Mettler Toledo, Denmark). The temperature was kept at 30 °C throughout the cultivation and pH maintained at pH 5.0 by automatic addition of 2 M NaOH and 2 M H_2SO_4. The bioreactor was sparged with sterile atmospheric air. Batch fermentation was run in duplicates. The cultivation medium had the following composition: 20 $g.L^{-1}$ sucrose; 1 $g.L^{-1}$ glucose; 10 $g.L^{-1}$ KNO_3; 8 $g.L^{-1}$ K_2PO_4; 1 $g.L^{-1}$ NaCl; 2 $g.L^{-1}$ $MgSO_4$; 0.5 $g.L^{-1}$ KCl; 0.1 $g.L^{-1}$ $CaCl{\cdot}2H_20$; 1 $mL.L^{-1}$ trace metal solution (composed of: 0.4 $g.L^{-1}$ $CuSO_4{\cdot}5H_2O$; 0.04 $g.L^{-1}$ $Na_2B_4O_7{\cdot}$ $10H_2O$; 0.8 $g.L^{-1}$ $FeSO_4{\cdot}$ $7H_2O$; 0.8 $g.L^{-1}$ $MnSO_4{\cdot}$ $2H_2O$; 0.8 $g.L^{-1}$ $Na_2MoO_4{\cdot}$ $2H_2O$; and 0.8 $g.L^{-1}$ $ZnSO_4{\cdot}$ $7H_2O$). The cultivations was inoculated with $1{\cdot}10^6$ $spores.L^{-1}$ and was run for six days after which the supernatant was collected by filtration, shielded against light and stored at 4°C until further processing

Fermentation broth from three reactor cultivations were extracted twice overnight with ethyl acetate (EtOAc) adjusted with formic acid to pH 3. The organic extracts were filtered and concentrated *in vacuo*. The target compound ($[M+H]^+$=541.2255) was enriched into a crude fraction by flash chromatography of the crude extract, using an Isolera One automated flash system (Biotage), using a water/methanol gradient elution. The final isolation of the compound was performed by a semi-preparative Waters 600 DAD on a LUNA-II column (250 mm x 10 mm, 5 μm) using a water/acetonitrile gradient (with TFA).

The LC-MS analyses were performed on a maXis quadrupole time of flight (qTOF) mass spectrometer (Bruker Daltonics) with an electrospray ionization (ESI) ion source. The maXis was calibrated using sodium formate automatically infused prior to each analytical run,

providing a mass accuracy of below 1 ppm. The mass spectrometer was linked to an Ultimate 3000 UHPLC system (Dionex) with DAD. Separation was achieved on a Kinetex C18, 2.6µm, 2.1x100 mm column (Phenomenex) with a flow of 0.4 ml min^{-1} at 40°C using a linear gradient 10 % acetonitrile (ACN) in Milli-Q water (MQ) with 20 µM formic acid (FA) going to 100 % ACN in 10 min.

1D and 2D NMR spectra (^{1}H, DQF-COSY, edHSQC, HMBC and NOESY) were recorded on either a Bruker Ascend 400 MHz (Bruker, Billerica, MA, USA) located at DTU, or on a Bruker Avance 800 MHz located at the Danish Instrument Centre for NMR spectroscopy of Biological Macromolecules at Carlsberg Laboratory. NMR spectra were acquired using standard pulse sequences. The solvent used was DMSO-d_6, which was also used as reference with signals at δ_H = 2.50 ppm and δ_C = 39.5 ppm. Data processing and analysis was done using MestReNova v.6.2.1-7569 (Mestrelab Research, Santiago de Compostela, Spain) and ACD NMR Workbook (Advanced Chemical Development, Inc., Toronto, Ontario, Canada). *J*-couplings are reported in hertz (Hz) and chemical shifts in ppm (δ).

B.3 Results

The purified N-glutamyl monascorubraminic acid was a dark red, almost black, amorphous solid. HR-ESI-MS gave a mass-to-charge ratio of $[M+H]^+$ = 541.2255, corresponding to a molecular formula of $C_{28}H_{32}N_2O_9$ (DBE=14).

The compound showed a UV absorption spectrum similar to that of known nitrogen containing *Monascus* pigments (Broder and Koehler, 1980; Hajjaj et al., 1997; Jung et al., 2003), with a UV_{max} at 516 nm (Figure B1).

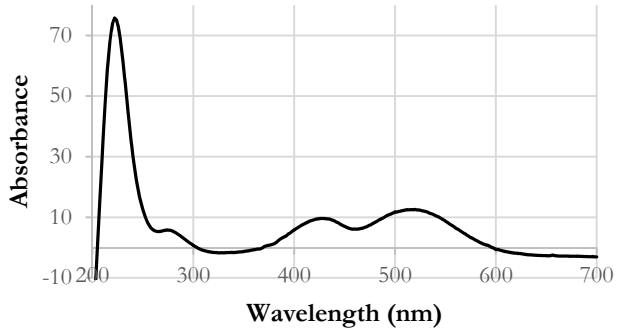


Figure B.1. UV/VIS spectra of N-glutamyl monascorubraminic acid

1D and 2D NMR (Table B1) were used to determine the structure of the compound (Figure B2). ^{1}H-NMR and HSQC revealed five olefinic protons [δ_H 8.22 ppm (s), δ_H 6.94 ppm (d, J = 11.8 Hz), δ_H 6.89 ppm (s), δ_H 6.59 ppm (s) and δ_H 6.37 ppm (d, J = 11.8 Hz)], two methyl [δ_H 1.59 ppm (s), and δ_H 0.85 ppm (t, J = 6.7 Hz)], eight CH_2 [δ_H 2.72 ppm (t, J=7.3 Hz), δ_H 2.33 ppm (m), δ_H 2.13 ppm (m), δ_H 1.48 ppm (quint, J=6.8), δ_H 1.24 ppm (m), and δ_H 1.23 ppm (a-c)(m)], and one methine [δ_H 4.93 ppm (m)].

^{13}C-NMR and HMBC revealed 12 quaternary carbons. Six carbonyls [δ_C 170.3, δ_C 194.6 ppm, δ_C 130.0 ppm, δ_C 172.9 ppm, δ_C 194.0 ppm, and δ_C 169.], five olefinic carbons [δ_C 117.2 ppm, δ_C 172.6 ppm, δ_C 149.4 ppm, δ_C 147.7 ppm, and δ_C 103.2 ppm], and one quaternary alkane [δ_C 84.9 ppm].

Additionally, two carboxylic acids [δ_H 12.5-14 (broad singlet) and δ_H 7.31 (s)] and an amide [δ_H 6.84 (s)] was present in the molecule.

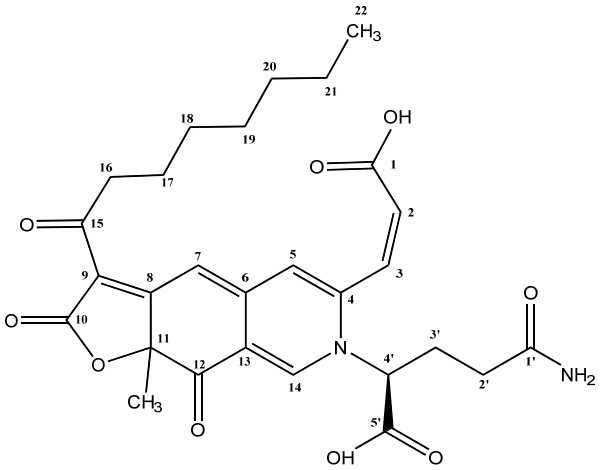


Figure B.2. Structure of the novel Monascus pigment N-glutamyl monascorubraminic acid.

Table B.1. Table showing proton and carbon shifts, multiplicity, as well as DQF-COSY and HMBC couplings.

#	1H	13C (HSQC)	Int.	Mult.	COSY	HMBC
1-OH	12.5-14	-	OH	bs	-	-
1	-	165.6	-	-	-	-
2	6.37	130.0	1H	d (*J*=11.8)	**3**	**1, 4, 7**
3	6.94	132.2	1H	d (*J*=11.8)	**2**	**1**
4	-	147.7	-	-	-	-
5	6.89	118.6	1H	s	-	**3, 4, 7, 13**
6	-	149.4	-	-	-	-
7	6.59	96.4	1H	s	-	**6, 8, 9, 11, 13, 15**
8	-	172.6	-	-	-	-
9	-	103.2	-	-	-	-
10	-	169.3	-	-	-	-
11	-	84.9	-	-	-	-
11-CH_3	1.59	29.6	3H	s	-	**8, 10, 11, 12**
12	-	194.0	-	-	-	-
13	-	117.2	-	-	-	-
14	8.22	140.5	1H	s	-	**6, 12, 13**
15	-	194.6	-	-	-	-
16	2.72	39.6	2H	t (*J*=7.3 Hz)		**15, 17, 18**
17	1.48	24.0	2H	quint (*J*=6.8 Hz)		**15, 16, 18**
18	1.23d	28.6	2H			
19	1.23a	28.9	2H	bs/m	**16-22**	**16-22**
20	1.23c	31.2	2H			
21	1.23b	22.1	2H			
22	0.85	13.9	3H	t (*J*=6.7 Hz)		**20, 21**
1'-NH2	6.84	-	NH2	s	**5'-OH**	-
1'	-	170.3	-	-	-	-
2'	2.13	30.6	2H	m	**3'**	**3', 5'**
3'	2.33	26.9	2H	m	**2', 4'**	**2', 5'**
4'	4.93	47.8	1H	m	**3'**	-
5'	-	172.9	-	-	-	-
5'-OH	7.31	-	OH	s	**1'-NH2**	-

By comparing the spectral data to literature on known *Monascus* pigments (Ogihara et al., 2000; Ogihara & Oishi, 2002), the obtained spectra indicated the compound to be a monascorubramine homologue, as also suggested by the UV spectrum.

DQF-COSY (Figure B3) showed 3J-couplings between 2' and 3', and 3' and 4', as well as between 2 and 3 (J=11.8, suggesting Z-configuration). Furthermore, the aliphatic chain consisting of 16-22 could be identified.

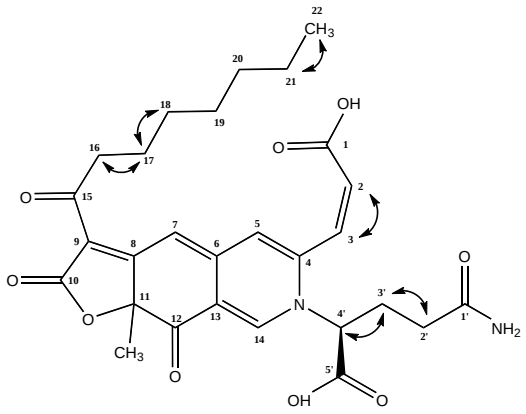


Figure B.3. Structure showing COSY-correlations

HMBC (Figure B4) provided long-range H-C-couplings within the azaphilone scaffold, linking **16** and **17** to the carbonyl **15**. **7** also showed a coupling to **15**, in addition to couplings to **8, 9**, **11**, **13** and **6**, while **14** had couplings to **12** and **13**. Protons **2** and **5** both exhibited couplings to **4**.

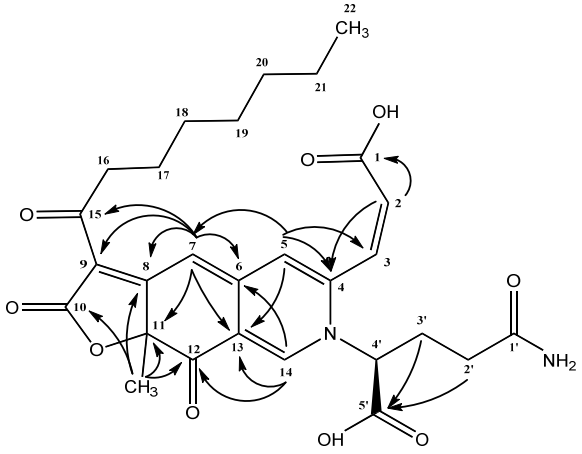


Figure B.4. Structure showing HMBC-correlations

NOESY experiments provided couplings between **2** and **3**, confirming the Z-relationship, as well as between **5** and **7**.

In conclusion, the compound is structurally similar to the monascorubramine PP-V discovered by Ogihara *et al.* (Ogihara et al., 2000), with the addition of a glutamine moiety to the isoquinoline system.

B.4 References

Hajjaj, H., Klaebe, A., Loret, M.O., Tzedakis, T., Goma, G., et al., **1997**. Production and Identification of N-Glucosylrubropunctamine and N-Glucosylmonascorubramine from Monascus ruber and Occurrence of Electron Donor-Acceptor Complexes in These Red Pigments. *Applied and environmental microbiology*, **63**: 2671–8.

Jung, H., Kim, C., Kim, K. & Shin, C.S., **2003**. Color characteristics of monascus pigments derived by fermentation with various amino acids. *Journal of agricultural and food chemistry*, **51**: 1302–6. doi:10.1021/jf0209387

Broder, C.U., Koehler, P.E., **1980**. Pigments produced by *Monascus purpureus* with regard to quality and quantity.*Journal of Food Science* , **45**: 567–569.

Ogihara, J., Kato, J., Oishi, K., Fujimoto, Y. & Eguchi, T., **2000**. Production and structural analysis of PP-V, a homologue of monascorubramine, produced by a new isolate of Penicillium sp. *Journal of bioscience and bioengineering*, **90**: 549–554.

Ogihara, J. & Oishi, K., **2002**. Effect of ammonium nitrate on the production of PP-V and monascorubrin homologues by Penicillium sp. AZ. *Journal of bioscience and bioengineering*, **93**: 54–9.

B.5 Supplementary material

Supplementary material contain NMR spectra of N-glutamyl monascorubraminic acid; ^{1}H, ^{13}C, DQF-COSY, NOESY, HSQC, HMBC.

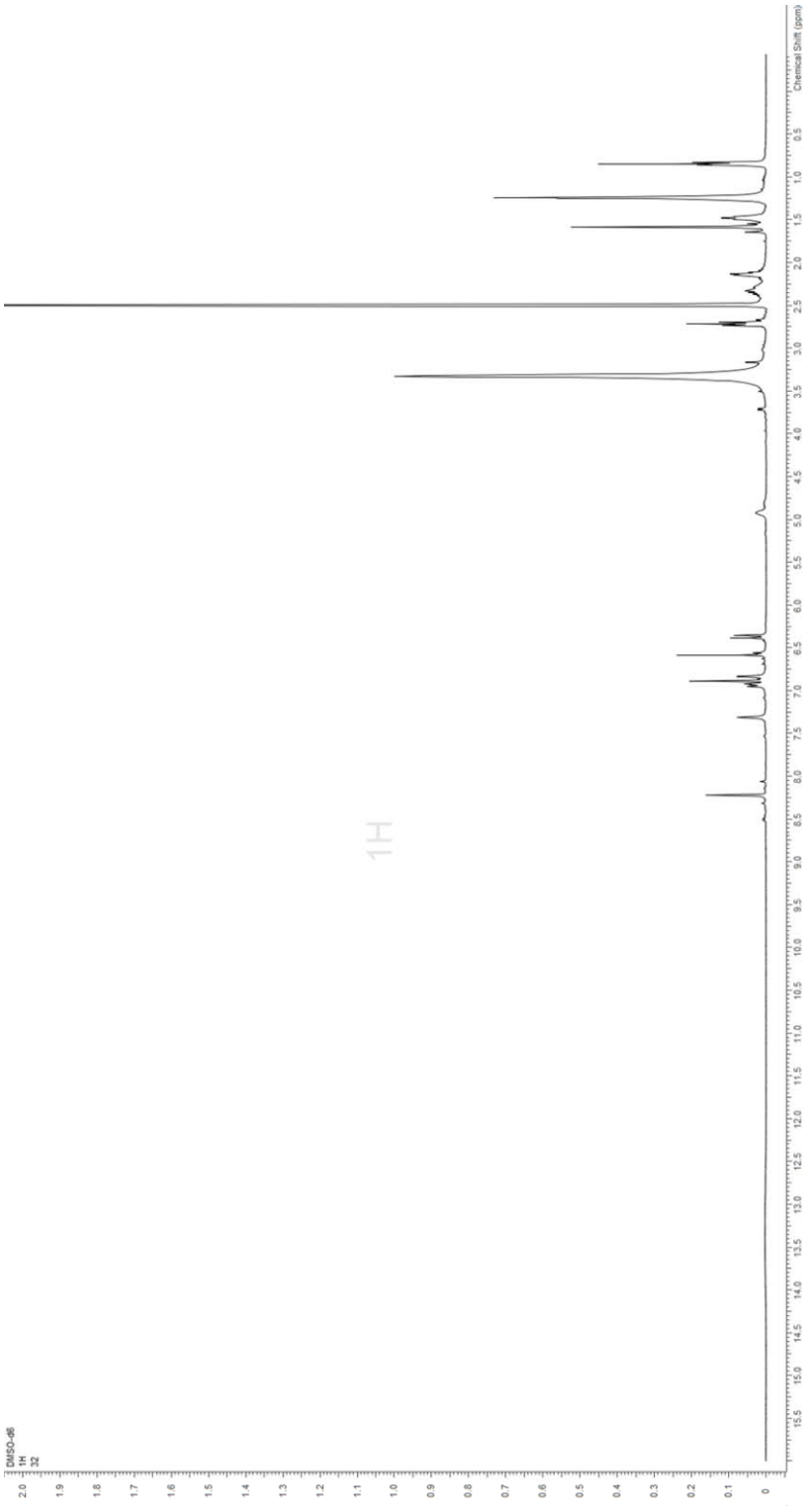


Figure B.5. 1H-NMR of N-glutamyl monascorubraminic acid

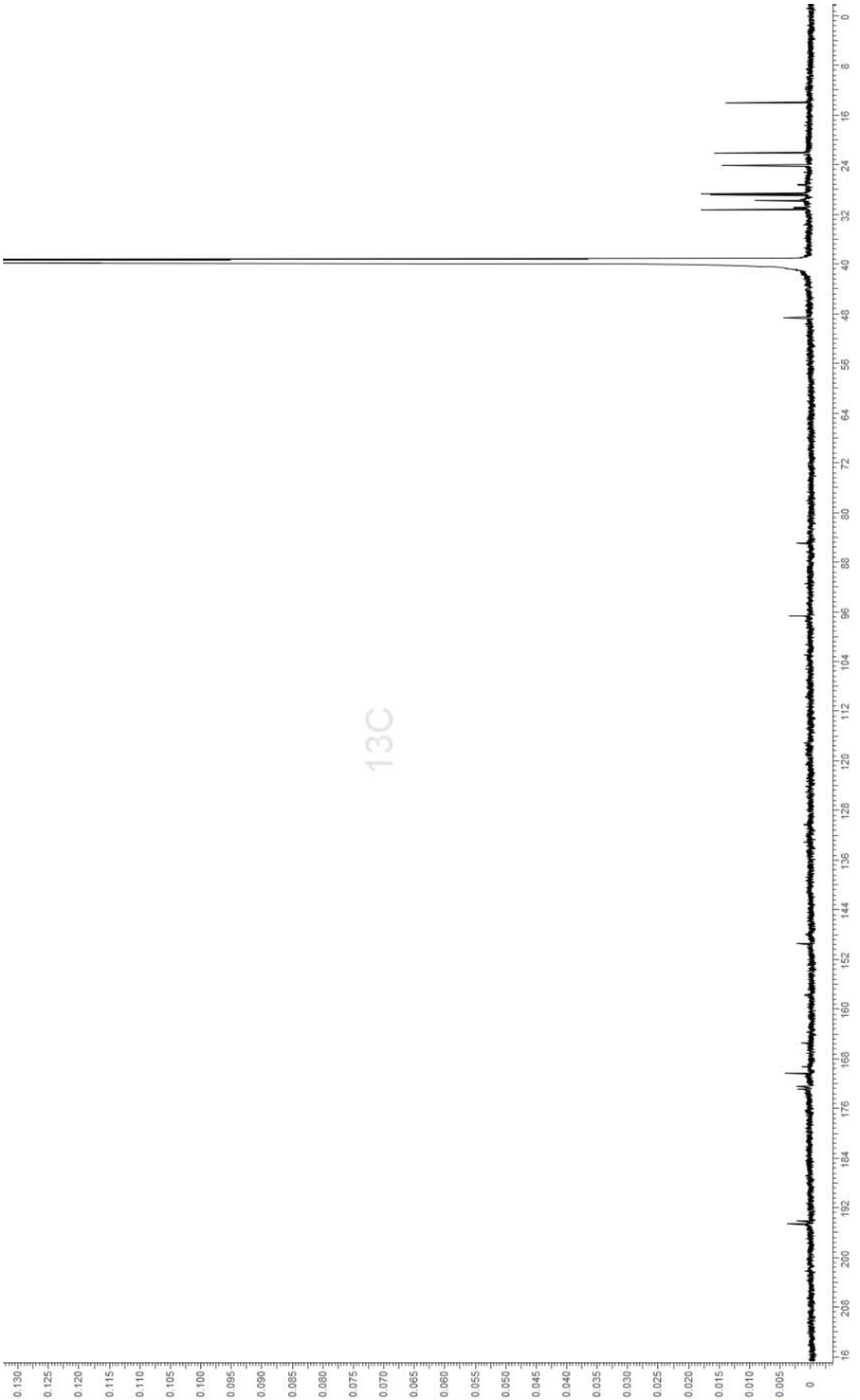


Figure B.6. 13C-NMR of N-glutamyl monascorubraminic acid

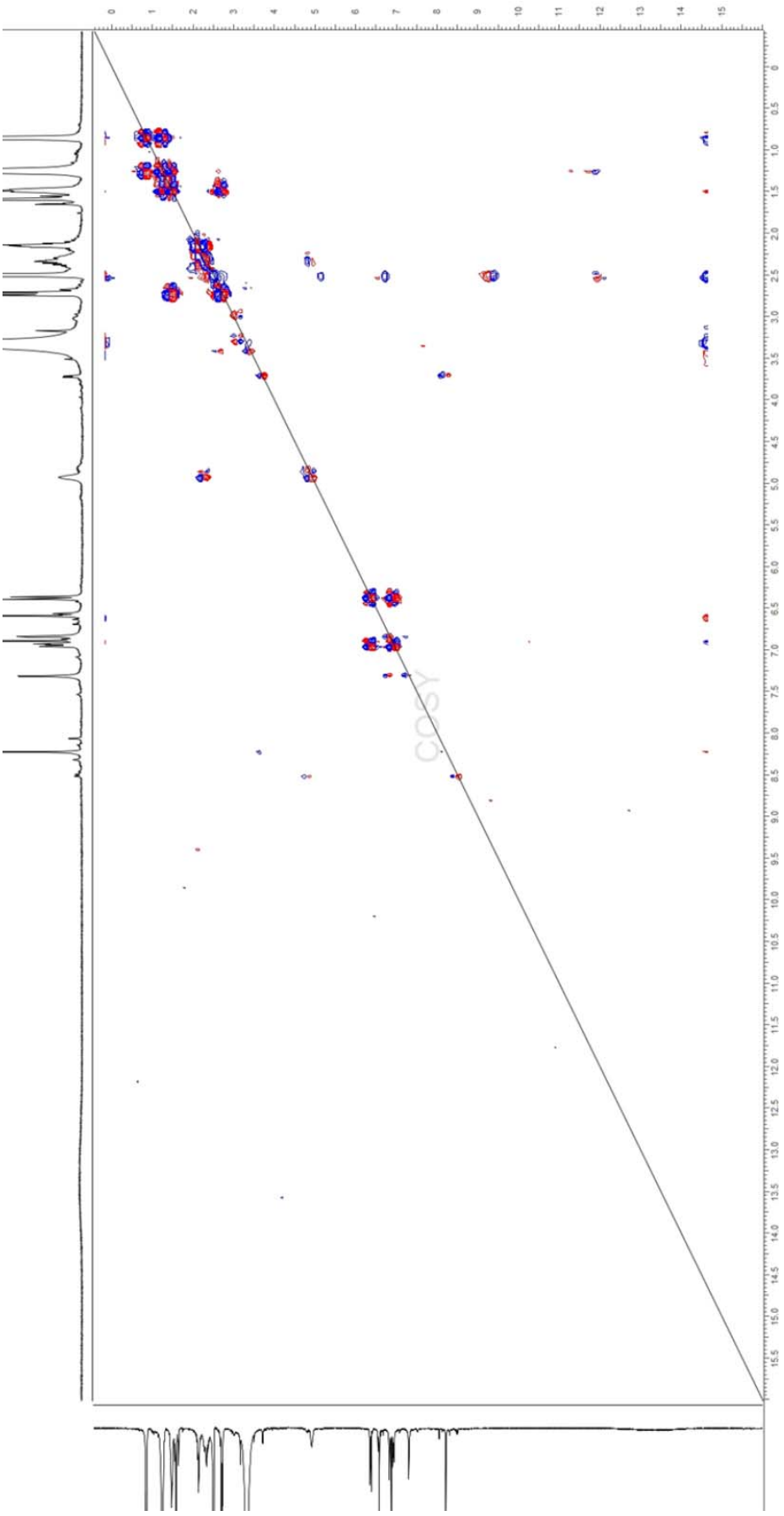


Figure B.7. DQF-COSY of N-glutamyl monascorubraminic acid

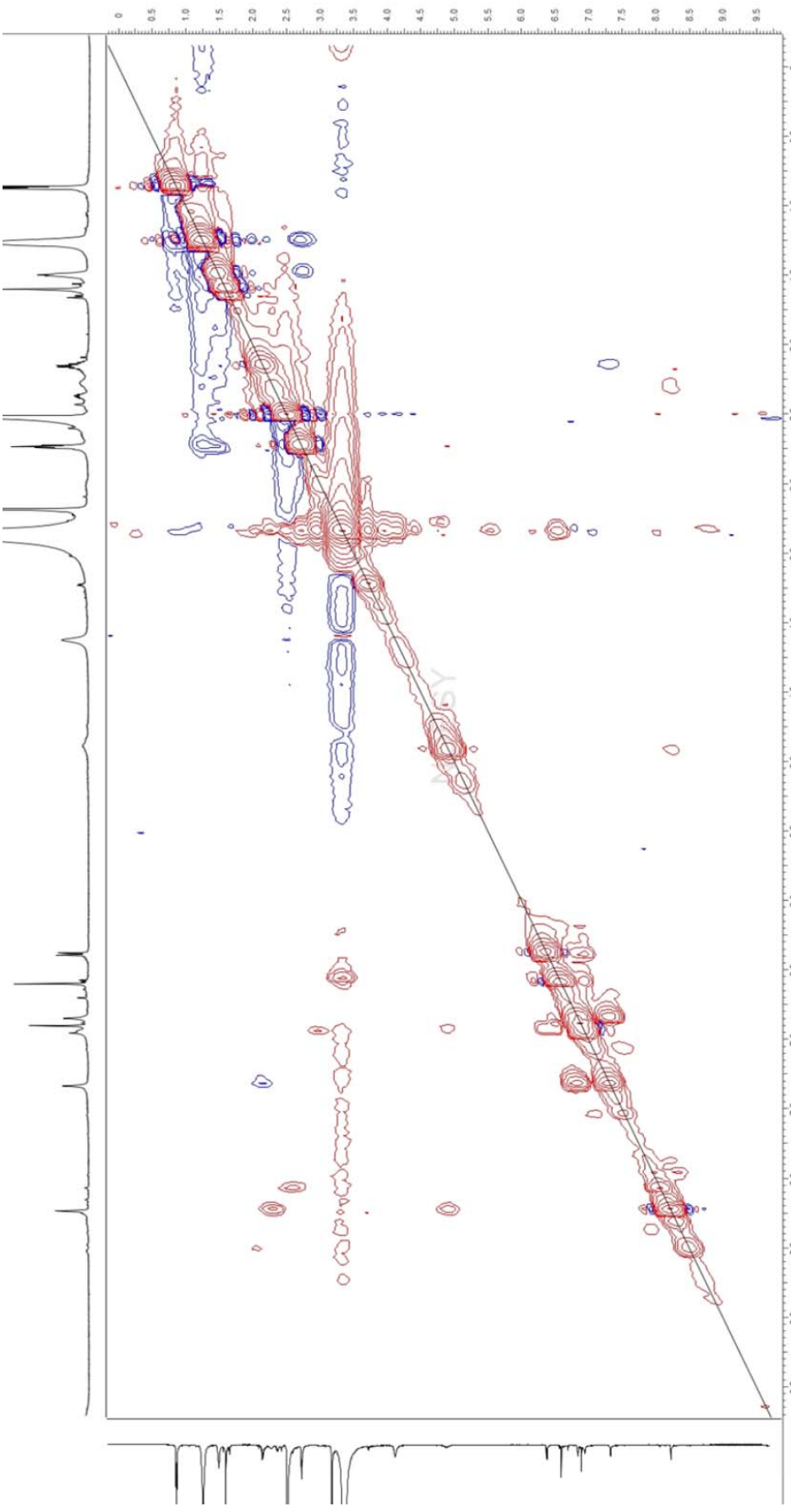


Figure B.8. NOESY of N-glutamyl monascorubraminic acid

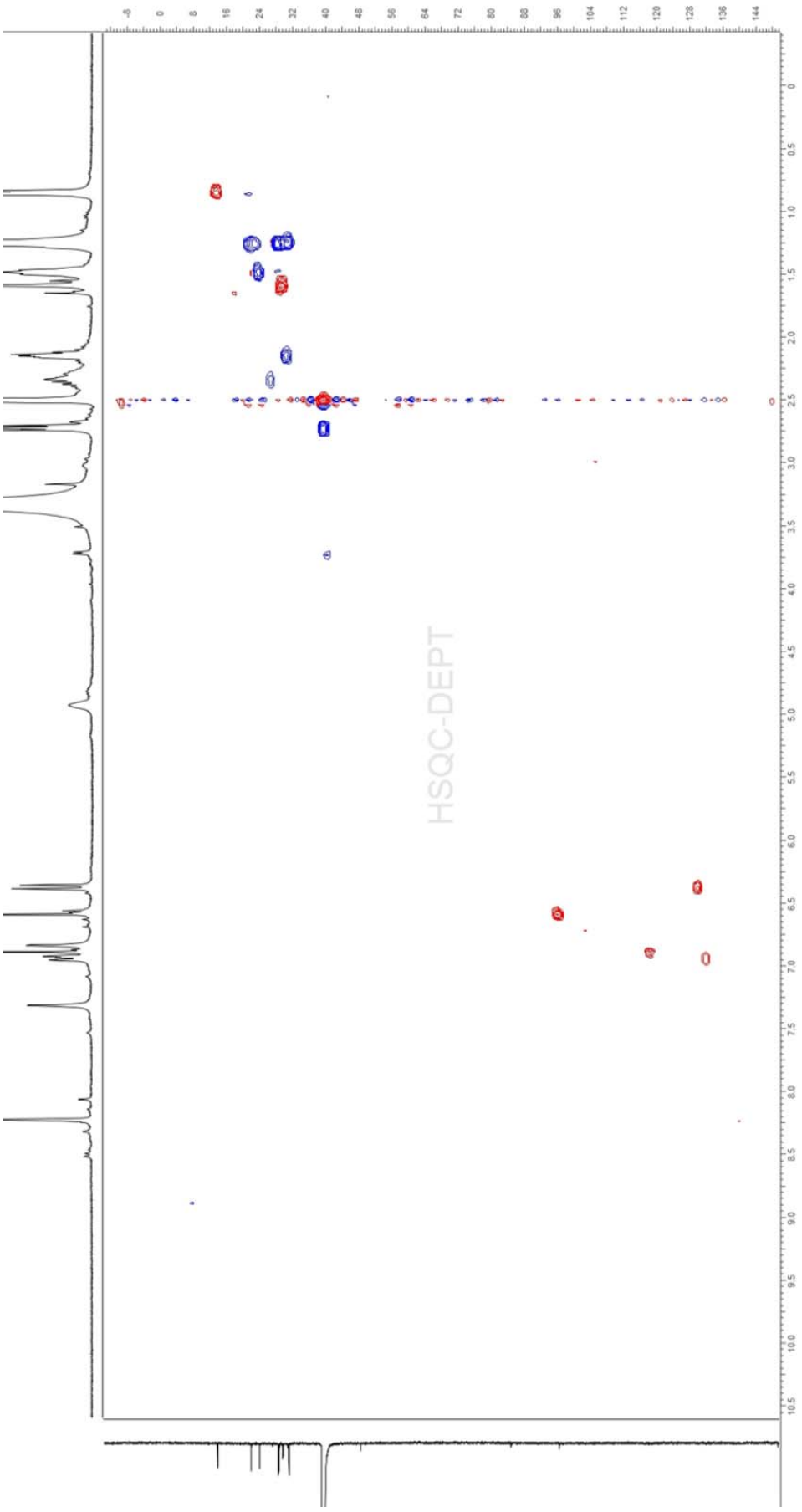


Figure B.9. HSQC of N-glutamyl monascorubraminic acid

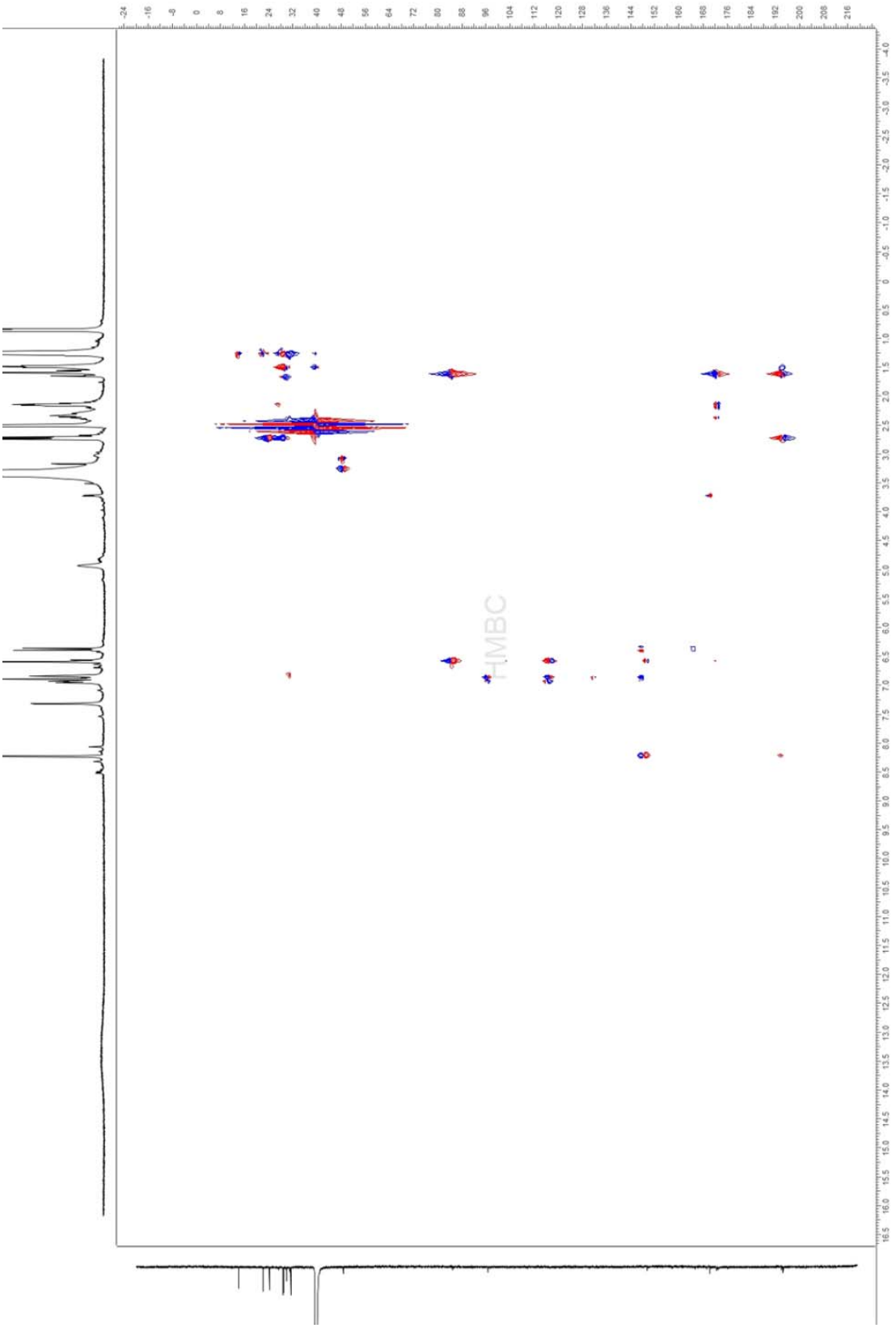


Figure B.10. HMBC of N-glutamyl monascorubraminic acid

Appendix C

Supporting information to Chapter 5

C.1 USER vector nomenclature

USER vectors follow the nomenclature introduced by (Hansen et al., 2011) and is used in the fungal USER plasmid collection at Section for Eukaryotic Biotechnology, Department of Systems Biology, Technical University of Denmark. Each USER vector is named in the following format:

pU QXYZ–IS–*YFG*

The U signifies that it is a USER compatible / USER created vector and *YFG* is Your Favourite Gene inserted in the vector. For the meaning of QXYZ and IS see the table below.

Table C.1. USER vector nomenclature as used in our laboratory

Q	**Marker**	**Note**
0	No marker	
1	*argB*	
2	AF*pyrG*	Flanked by direct repeats to allow excision
3	*ble*	Resistance to bleomycin and related genotoxins
4	*hyg*	Resistance to hygromycin
5	*trpC*	
6	*ble-TK*	Flanked by direct repeats to allow excision
X	**Promoter**	
0	No promoter	
1	P*gpdA*	For constitutive expression
2	P*alcA*	For constitutive expression
3	Tet-ON	For inducible expression
Y	**Terminator**	
0	No terminator	
1	T*trpC*	
Z	**USER Casettes**	
0	No casettes	
1	AsiSI/Nb.BtsI	
2	PacI/Nt.BbvC1	PacI/Nt.BbvC1, cassettes A and B
3	AsiSI/Nb.BtsI + PacI/Nt.BbvC1	PacI/Nt.BbvC1, cassettes A and B
4	AsiSI/Nb.BsmI + PacI/Nt.BbvC1	PacI/Nt.BbvC1, cassettes A and B
5	PacI/Nt.BtsI	Cassette A for gene insertion
6	AsiSI/Nb.BsmI	
IS	**Integration Site**	
0	No Integration site	
1	IS1	*A. nidulans*
2	IS2	*A. nidulans*
3	IS3	*A. nidulans*
4	IS4	*A. nidulans*
5	IS5	*A. nidulans*
R	AMA1-plasmid	Replicative plasmid

C.2 PKS genes in the genus of *Talaromyces* and their orthology

PKS domain abbreviations: KS: β-ketosynthase, AT: acyl trasnferase, PT: Product template, ACP: Acyl carrier protein, DH: Dehydrogenase, ER: enoyl reductase, KR: ketoreductase, cMT: C-methyl transferase, C: condensation domain, A: adenylation domain, TE: Thioesterase, R: Thioester reductase, PO: prolyl oligopeptidase

Table C.2. PKS genes in *Talaromyces marneffei*

PM1	ATCC18224	Domain architecture	Notes
GQ26_0093000	PMAA_001920	SAT-KS-AT-PT-ACP-ACP-TE	
GQ26_0180530	PMAA_100360	SAT-KS-AT-PT-ACP	
GQ26_0181690	PMAA_101550	SAT-KS-AT-PT-ACP-cMT-R	Mitorubrins
GQ26_0181740	PMAA_101600	SAT-KS-AT-PT-ACP-ACP	Mitorubrins
GQ26_0680070	PMAA_008670	KS-AT-DH-ER-KR-ACP	
GQ26_0032550	PMAA_031640	SAT-KS-AT-PT-ACP	
GQ26_0260910	PMAA_063620	SAT-KS-AT-PT-ACP-ACP-TE	
GQ26_0570080	PMAA_061720	SAT-KS-AT-PT-ACP-TE	
GQ26_0570130	PMAA_061670	KS-AT-DH-ER-KR-ACP	
GQ26_0470260	PMAA_095560	SAT-KS-AT-PT-ACP-ACP-TE	
GQ26_0260090	PMAA_045640	KS-AT-DH-cMT-ER-KR-ACP-CT	
GQ26_0041870	PMAA_088150	KS-AT-DH-cMT-KR-ACP-C-A-ACP-R	
GQ26_0310840	PMAA_066670	SAT-KS-AT-PT-ACP-TE	
GQ26_0310890	PMAA_066720	KS-AT-DH-ER-KR-ACP	
GQ26_0580350	PMAA_062940	SAT-KS-AT-PT-ACP-ACP-cMT	
GQ26_0022580	PMAA_009360	KS-AT-DH-ER-KR-ACP	
GQ26_0610090	PMAA_001080	KS-AT-DH-cMT-ER-KR-ACP	
GQ26_0022560	PMAA_009380	KS-AT-DH-cMT-KR-ACP-R	
GQ26_0231610	PMAA_101860	SAT-KS-AT-PT-ACP-ACP-cMT-R	*Monascus* pigments
pks4[a]	PMAA_082120	SAT-KS-AT-PT-ACP-ACP-TE	Melanin
pks5[a]	PMAA_081530	KS-AT-DH-ER-KR-ACP-C	
GQ26_0330880	PMAA_068360	KS-AT-DH-cMT-ER-KR-ACP	
GQ26_0120340	PMAA_039120	KS-AT-DH-ER-KR-ACP	
GQ26_0170630	PMAA_077910	KS-AT-DH-cMT-KR-ACP-C-A-ACP-R	
GQ26_0650100	PMAA_050380	KS-AT-DH-cMT-KR-ACP	
GQ26_0500060[b]	n/a	SAT-KS-AT-PT-ACP-ACP-cMT	

[a] PKS genes not found in the most recent assembly of *T. marneffei* PM1. The PKS genes however were assembled in the original PM1 genome assembly relying only on Sanger sequencing.
[b] PKS gene only present in the most recent assembly of *T. marneffei* PM1.
n/a: not available. This PKS gene is not detected in *T. marneffei* ATCC18224

Table C.3. PKS genes in *Talaromyces stipitatus* ATCC10500

Locus tag	Domain architecture	Known products
TSTA_000310	KS-AT-DH-cMT-ER-KR-ACP	
TSTA_001780	SAT-KS-AT-PT-ACP-cMT	
TSTA_001920	SAT-KS-AT-PT-ACP-cMT-R	
TSTA_004020	KS-AT-DH-cMT-ER-KR-ACP	
TSTA_008130	KS-AT-DH-cMT-ER-KR-ACP	
TSTA_008140	SAT-KS-AT-PT-ACP-ACP-TE	
TSTA_008950	SAT-KS-AT-PT-ACP-TE	
TSTA_009240	KS-AT-DH-ER-KR-ACP	
TSTA_014510	SAT-KS-AT-PT-ACP-ACP-TE	
TSTA_017100	KS-AT-DH-ER-KR-ACP	
TSTA_017120	KS-AT-DH-cMT-KR-ACP-R	
TSTA_030950	KS-AT-DH-ER-KR-ACP	
TSTA_048430	KS-AT-DH-cMT-ER-KR-ACP	
TSTA_049850	KS-AT-DH-ER-KR-ACP	
TSTA_051450	KS-AT-DH-cMT-ER-KR-ACP	
TSTA_052680	SAT-KS-AT-PT-ACP-ACP-TE	
TSTA_055600	KS-AT-DH-cMT-ER-KR-ACP	
TSTA_060720	SAT-KS-AT-PT-ACP-ACP-cMT-TE	
TSTA_061410	SAT-KS-AT-PT-ACP-ACP-cMT-R	*Monascus* pigments
TSTA_061660	SAT-KS-AT-PT-ACP-ACP	Mitorubrins
TSTA_061710	SAT-KS-AT-PT-ACP-cMT-R	Mitorubrins
TSTA_079690	KS-AT-DH-ER-KR-ACP	
TSTA_082870	SAT-KS-AT-PT-ACP-ACP-TE	
TSTA_083310	SAT-KS-AT-PT-ACP-ACP-TE	
TSTA_084970	KS-AT-DH-ER-KR	
TSTA_088080	KS-AT-KR-ACP-C-A-ACP-R	
TSTA_098280	KS-AT-DH-cMT-ER-KR-ACP	
TSTA_099020	KS-AT-DH-cMT-ER-KR-ACP	
TSTA_100460 + TSTA_100450[a]	SAT-KS-AT-PT	
TSTA_104650	KS-AT-DH-cMT-ER-KR-ACP	
TSTA_105770	KS-AT-DH-cMT-KR-ACP-C-A-ACP-R	
TSTA_106020	KS-AT-DH-cMT-ER-KR-ACP	
TSTA_106730	KS-AT-DH-cMT-ER-KR-ACP-C-A-ACP-R	
TSTA_116790	SAT-KS-AT-PT-ACP-ACP-TE	Melanin
TSTA_117750	SAT-KS-AT-PT-ACP-ACP-cMT-R	Tropolones, Stipitatic acid
TSTA_118920	KS-AT-DH-ER-KR-ACP	
TSTA_125550	KS-AT-DH-ER-KR-ACP	
TSTA_126190	SAT-KS-AT-PT-ACP	
TSTA_126840	SAT-KS-AT-PT-ACP	

[a]Gene calling with Augustus show that TSTA_100460 and TSTA_100450 constitute one PKS encoding gene

Table C.4. PKS genes in *Talaromyces aculeatus* ATCC10409

	Domain architecture	Known products
362170	KS-AT-DH-cMT-ER-KR-ACP	
367181	KS-AT-DH-cMT-ER-KR-ACP	
367683	KS-AT-DH-cMT-ER-KR-ACP	
376185	SAT-KS-AT-PT-ACP-ACP-TE	
377768	KS-AT-DH-ER-ACP-CT	
385053	KS-AT-DH-cMT-ER-KR-ACP	
386035	SAT-KS-AT-PT-ACP-ACP-cMT-R	*Monascus* pigments
404794	SAT-KS-AT-PT-ACP	
404907	SAT-KS-AT-PT-ACP-ACP-R	
405069	SAT-KS-AT-PT-ACP-cMT-R	Mitorubrins
405085	SAT-KS-AT-PT-ACP-ACP	Mitorubrins
407895	SAT-KS-AT-PT-ACP-cMT-R	
410681	SAT-KS-AT-PT-ACP-ACP-TE	
414781	KS-AT-DH-cMT-ER-KR-ACP	
416266	KS-AT-DH-ER-KR-ACP	
416297	KS-AT-DH-cMT-ER-KR-ACP	
419872	SAT-KS-AT-PT-ACP-TE	
431673	KS-AT-DH-ER-KR-ACP	
431674	SAT-KS-AT-PT-ACP-ACP-TE	
446735	KS-AT-DH-cMT-ER-KR-ACP	
451584	KS-AT-DH-ER-KR-ACP	
455096	KS-AT-DH-cMT-ER-KR-ACP	
456252	KS-AT-DH-cMT-ER-KR-ACP	
456509	SAT-KS-AT-PT-ACP-cMT-TE	
457238	KS-AT-DH-cMT-ER-KR-ACP	
460748	KS-AT-DH-ER-KR-ACP	
468088	SAT-KS-AT-PT-ACP-cMT-R	
471040	KS-AT-DH-cMT-ER-KR-ACP	
477924	SAT-KS-AT-PT-ACP-ACP-TE	
482472	KS-AT-DH-cMT-ER-KR-ACP	
495414	KS-AT-DH-ER-KR-ACP	
501053	KS-AT-DH-cMT- -ACP-C-A-ACP-R	
501902	SAT-KS-AT-PT-ACP-TE	
502117	KS-AT-DH-cMT-ER-KR-ACP	
503264	KS-AT-DH-cMT-ER-KR-ACP	
509739	SAT-KS-AT-PT-ACP	

Table C.5. PKS genes in *Talaromyces atroroseus* IBT11181

locus tag	Domain architecture	Known products
UA08_00425	SAT-KS-AT-PT-ACP-ACP-TE	Melanin
UA08_00474	KS-AT-DH-KR-ACP-C-A-ACP-R	
UA08_00993	SAT-KS-AT-PT-ACP-ACP-cMT-R	
UA08_01172	SAT-KS-AT-PT-ACP	
UA08_01303	KS-AT-DH-ER-KR-ACP-C-A-ACP-TE	
UA08_02259	KS-AT-DH-ER-KR-ACP	
UA08_02873	KS-AT-DH-cMT-ER-KR-ACP	
UA08_02969	KS-AT-DH-cMT-KR-ACP-TE	
UA08_03198	KS-AT-DH-cMT-KR-ACP-C-A-ACP-TE	
UA08_04451	KS-AT-DH-cMT-KR-ACP-C-A-ACP-TE	
UA08_04929	KS-AT-DH-ER-KR-ACP	
UA08_04992	SAT-KS-AT-PT-ACP-cMT-PO/TE	
UA08_05018	KS-AT-DH-cMT-KR-ACP-C-A-ACP-TE	
UA08_05979	SAT-KS-AT-PT-ACP-ACP	
UA08_05984	SAT-KS-AT-PT-ACP-cMT-R	Mitorubrins
UA08_06759	KS-AT-DH-cMT-ER-KR-ACP	Mitorubrins
UA08_06889	KS-AT-DH-cMT	
UA08_07371	KS-AT-DH-cMT-ER-KR-ACP	
UA08_07744	KS-AT-DH-ER-KR-ACP	
UA08_09373	SAT-KS-AT-PT-ACP	

Table C.6. Orthologues PKS genes in genome sequenced *Talaromyces* species. The PKS genes are identified by their locus tags, except for *T. aculeatus* PKS genes, which are identified by their transcript_id retrieved from the JGI genome portal.

T. marneffei **ATCC18224**	***T. stipitatus*** **ATCC10500**	***T. aculeatus*** **ATCC10409**	***T. atroroseus*** **IBT11181**
PMAA_001920	TSTA_014510	376182	
PMAA_101860	TSTA_061410	386035	
PMAA_082120	TSTA_116790		UA08_00425
PMAA_068360		362170	
PMAA_039120	TSTA_079690	416266	UA08_04929
PMAA_100360		404794	
PMAA_101550	TSTA_061710	405069	UA08_5984
PMAA_101600	TSTA_061660	405085	UA08_5979
PMAA_008670	TSTA_030950		
PMAA_063620	TSTA_083310		
PMAA_066670		419872	
PMAA_066720		460748	
PMAA_009360	TSTA_017100		
PMAA_001080	TSTA_098280	414781	
PMAA_009380	TSTA_017120		
	TSTA_004020	385053	
	TSTA_049850	495414	
	TSTA_100460		UA08_09374
	TSTA_104650	482472	
		501053	UA08_03198
	TSTA_106020		UA08_02259

C.3 Chemical analysis of produced pigments in *T. atroroseus*

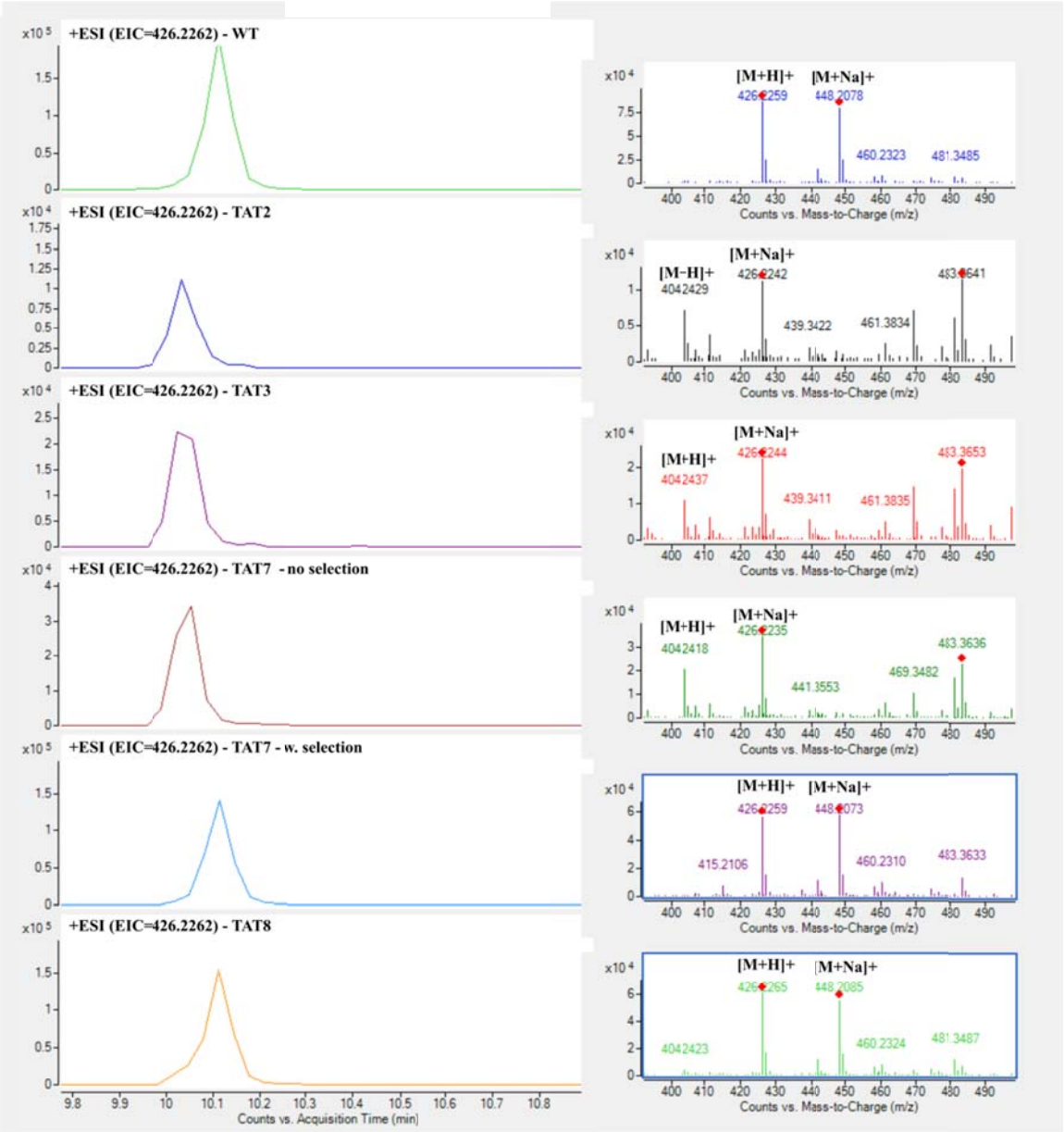


Figure C.1. EIC and MS spectra of m/z 426.2262 relating to compound 6 (PP-R). On the left the peak in the EIC of TAT2, TAT3 and TAT7 (first group) is shifted in time in relation to the peaks of WT, TAT7 w. selection and TAT8 (second group). Furthermore the height of the peaks is significant lower in the first group. On the right the corresponding MS spectra show that the accurate mass in the first group is slightly lower than in the second group, and that the m/z around 426.2242 in the first groups is actually the sodium adduct relating to the $[M+H]^+$ around 404.2429.

Appendix D

Supporting information to Chapter 6

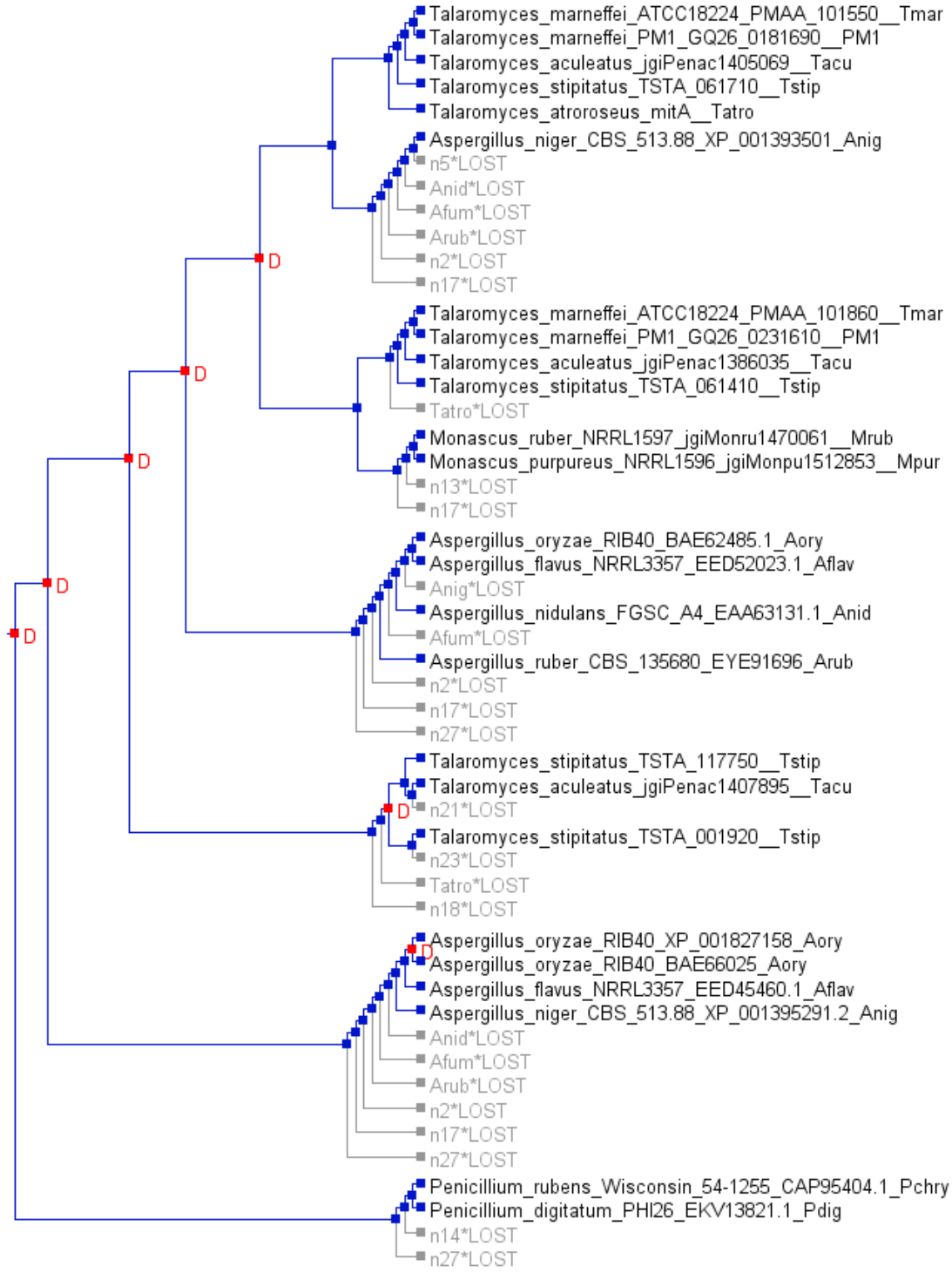


Figure D.1. Reconciled MitA gene tree using Notung v2.8. All combinations of DTL costs applied that produced temporal feasible solutions ended with the same reconciliation. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches.

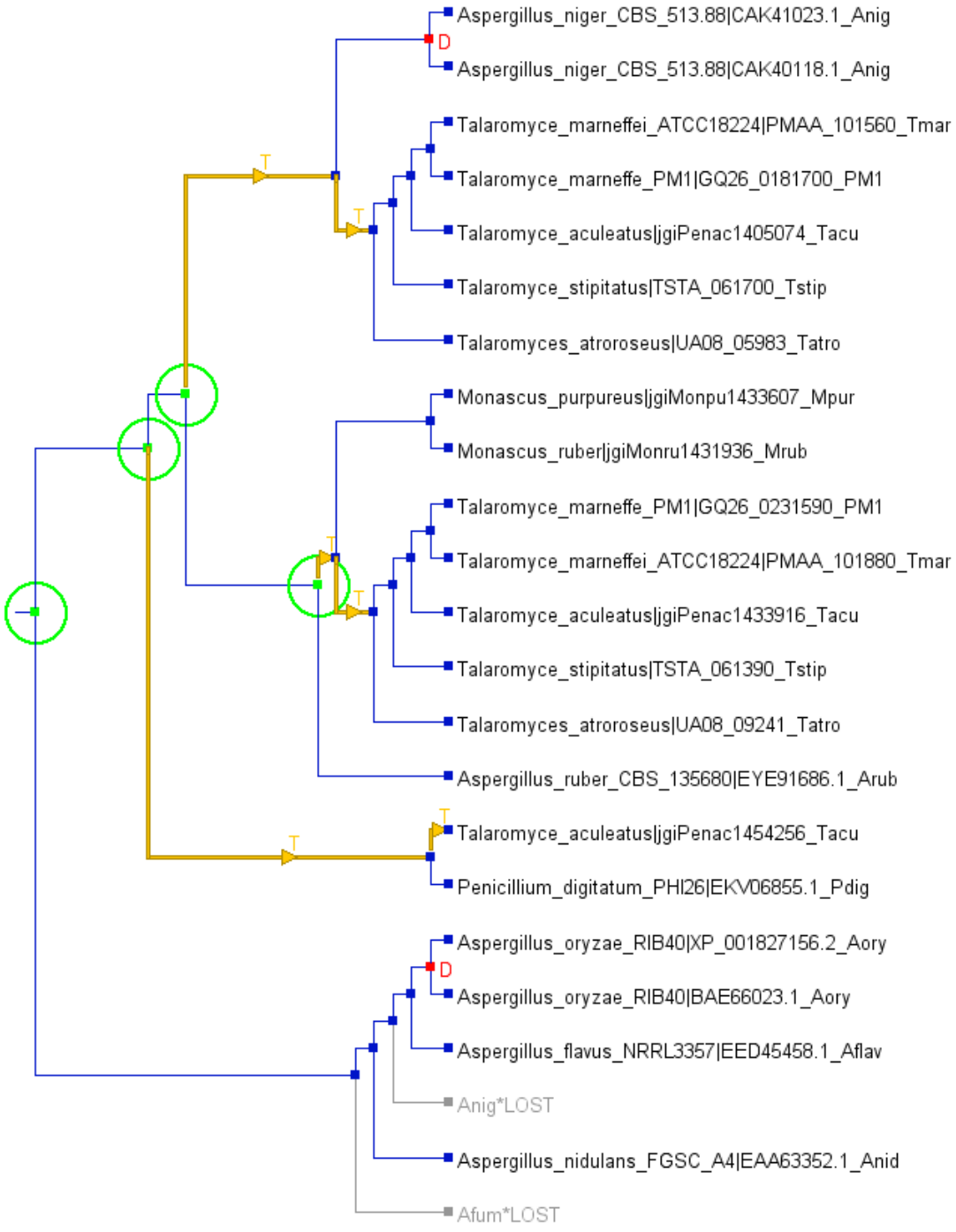


Figure D.2. Reconciled UA08_05983 gene tree #1 using Notung v2.8. Maximum parsimonious reconciliation tree #1 for UA08_05983 using duplication cost of 2, transfer cost of 3 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. These solutions are shown in UA08_05983 gene trees #2-4 shown in Figure F3-5.

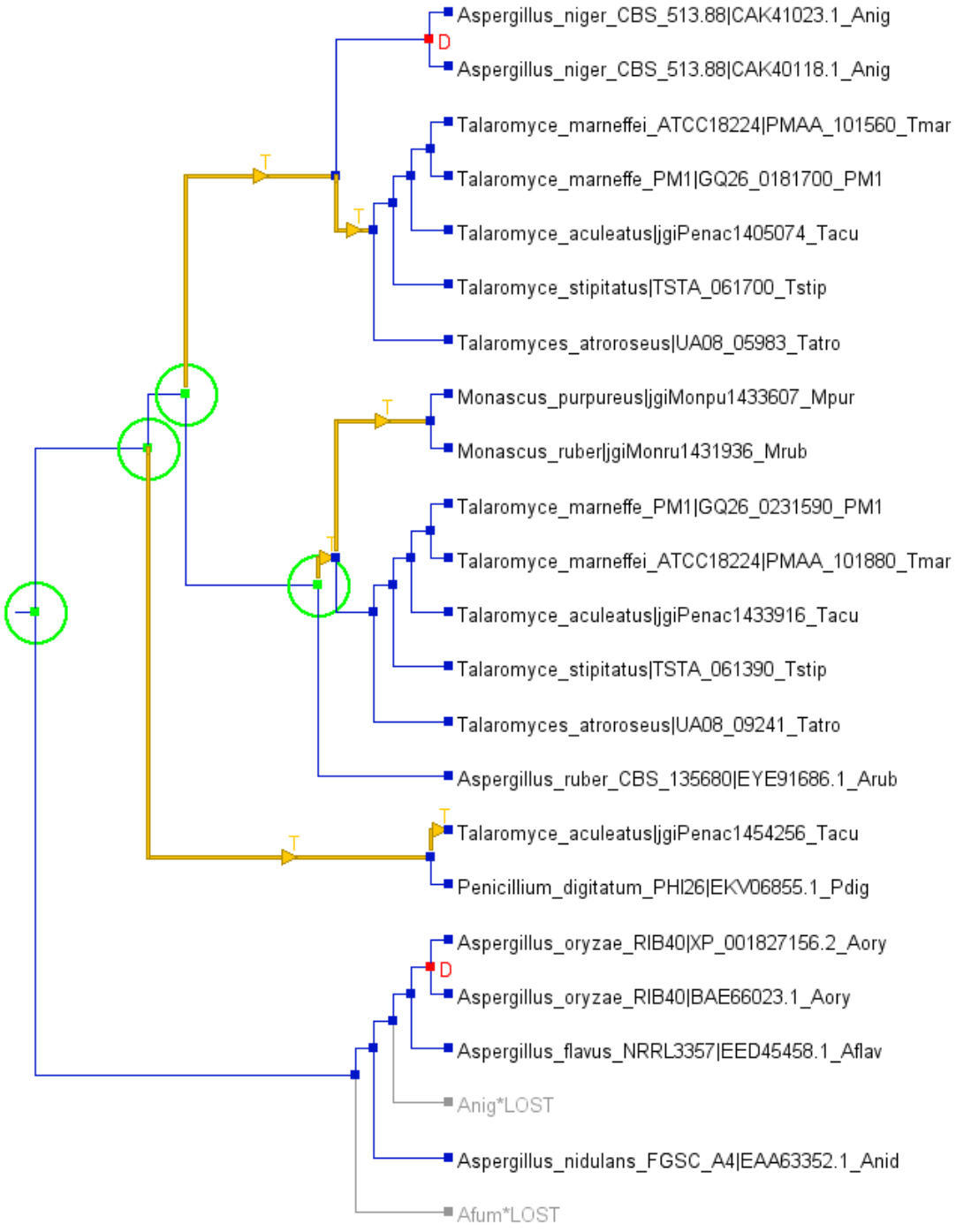


Figure D.3. Reconciled UA08_05983 gene tree #2 using Notung v2.8. Maximum parsimonious reconciliation tree #2 for UA08_05983 using duplication cost of 2, transfer cost of 3 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. These solutions are shown in UA08_05983 gene trees #1,#3-4 shown in Figure F2,F4-5.

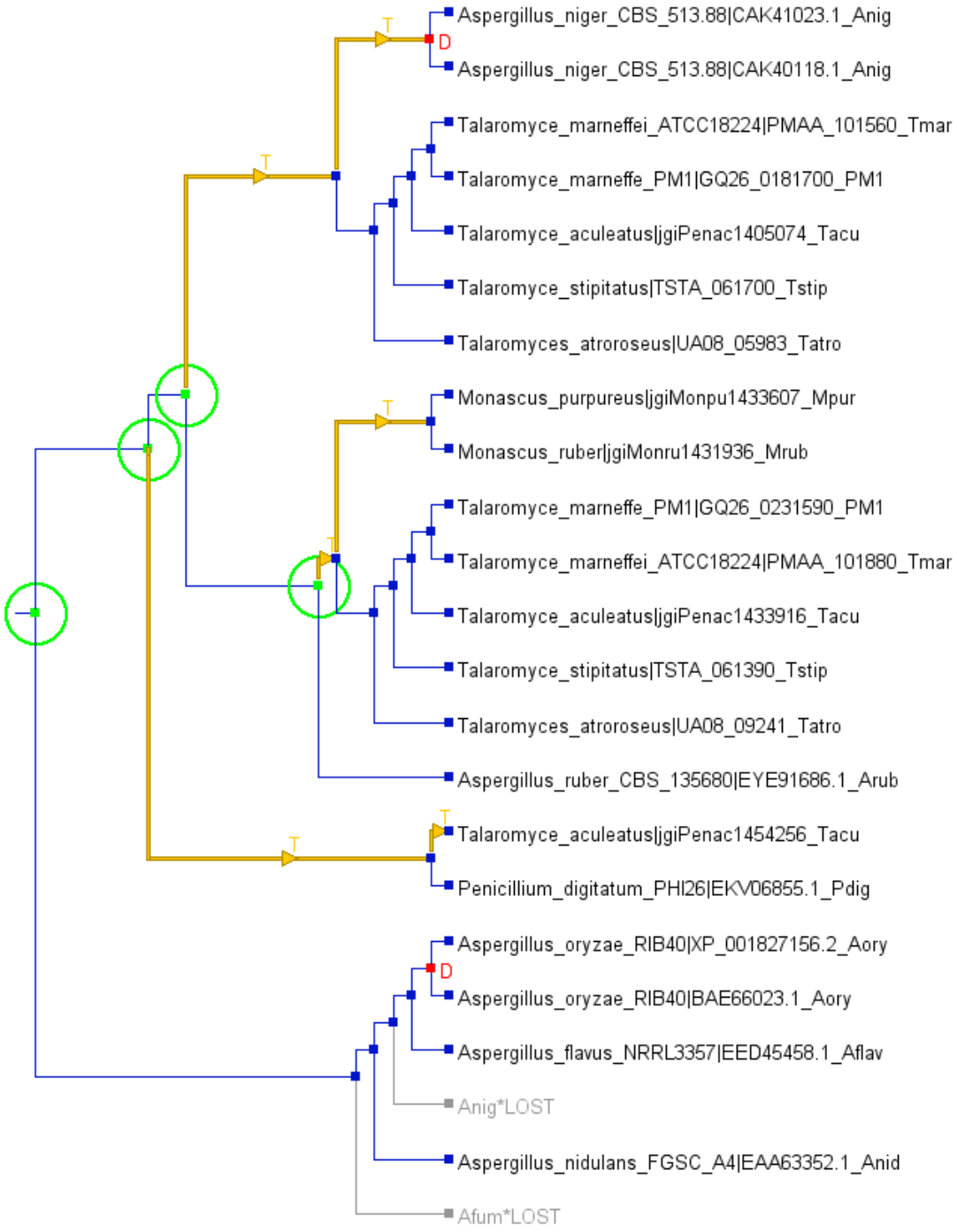


Figure D.4. Reconciled UA08_05983 gene tree #3 using Notung v2.8. Maximum parsimonious reconciliation tree #3 for UA08_05983 using duplication cost of 2, transfer cost of 3 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. These solutions are shown in UA08_05983 gene trees #1-2,#4 shown in Figure F2-3,F5.

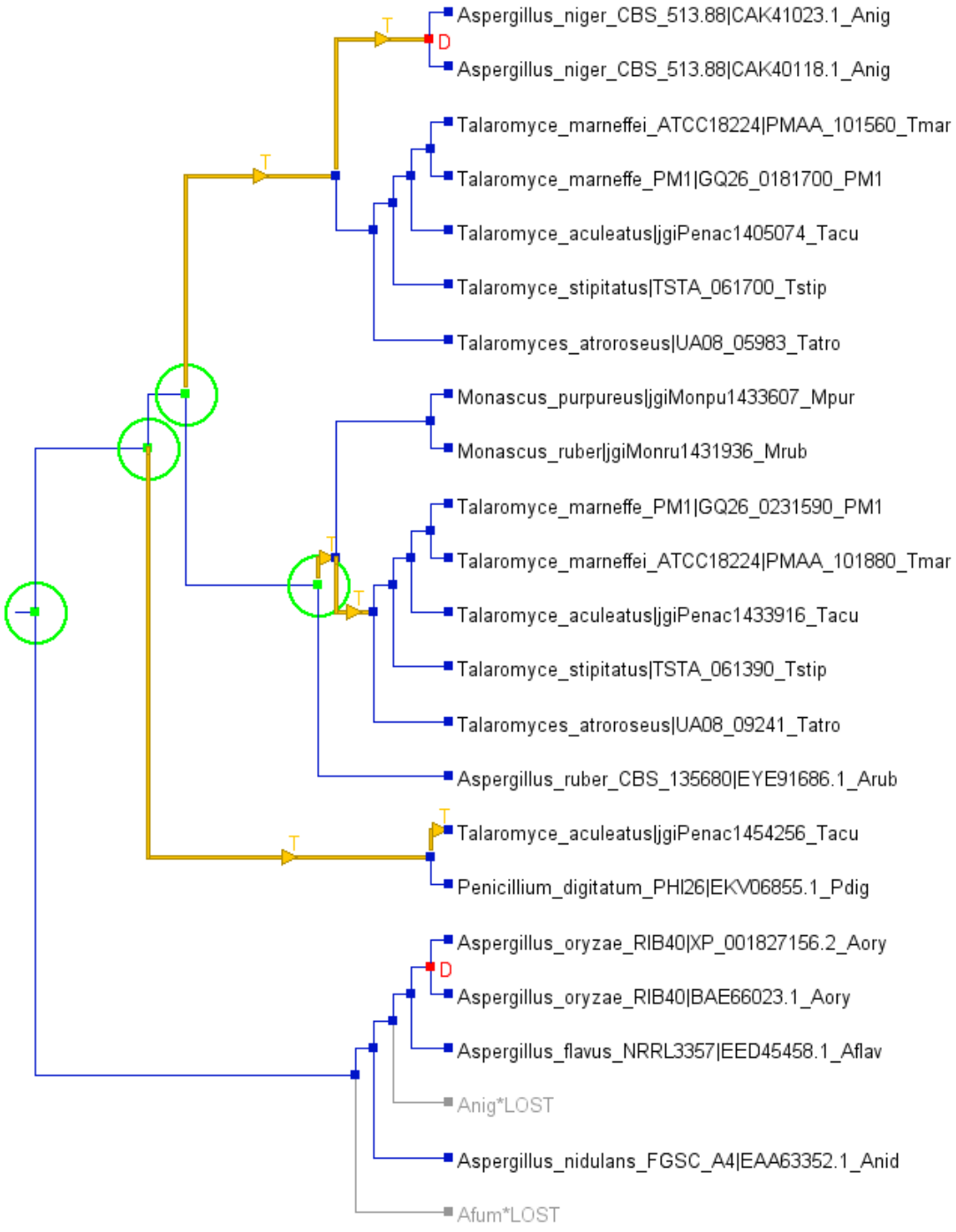


Figure D.5. Reconciled UA08_05983 gene tree #4 using Notung v2.8. Maximum parsimonious reconciliation tree #4 for UA08_05983 using duplication cost of 2, transfer cost of 3 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. These solutions are shown in UA08_05983 gene trees #1-3 shown in Figure F2-4.

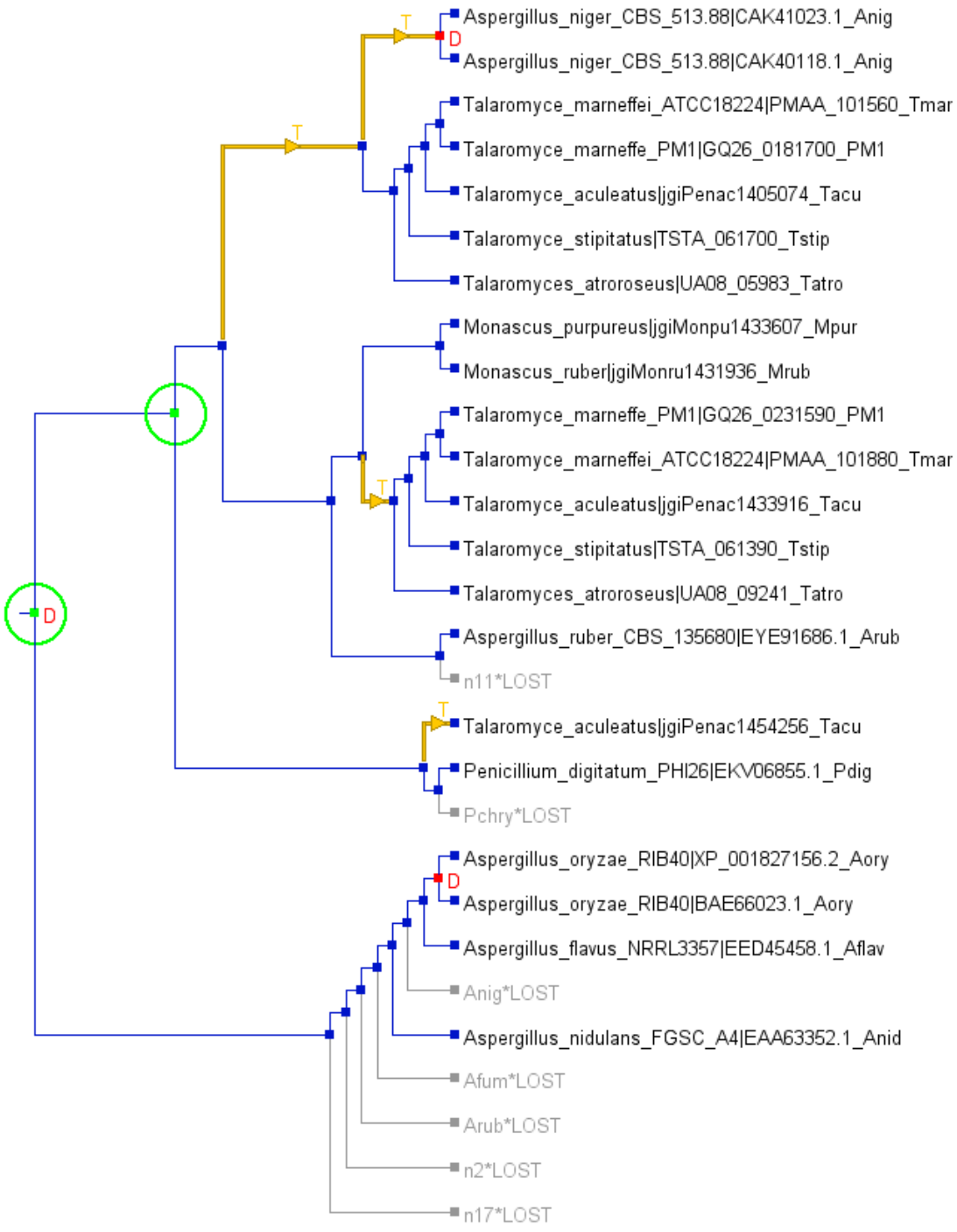


Figure D.6. Reconciled UA08_05983 gene tree #5 using Notung v2.8. Maximum parsimonious reconciliation tree #5 for UA08_05983 using duplication cost of 2, transfer cost of 4 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. These solutions are shown in UA08_05983 gene trees #6-7 shown in Figure F7-8.

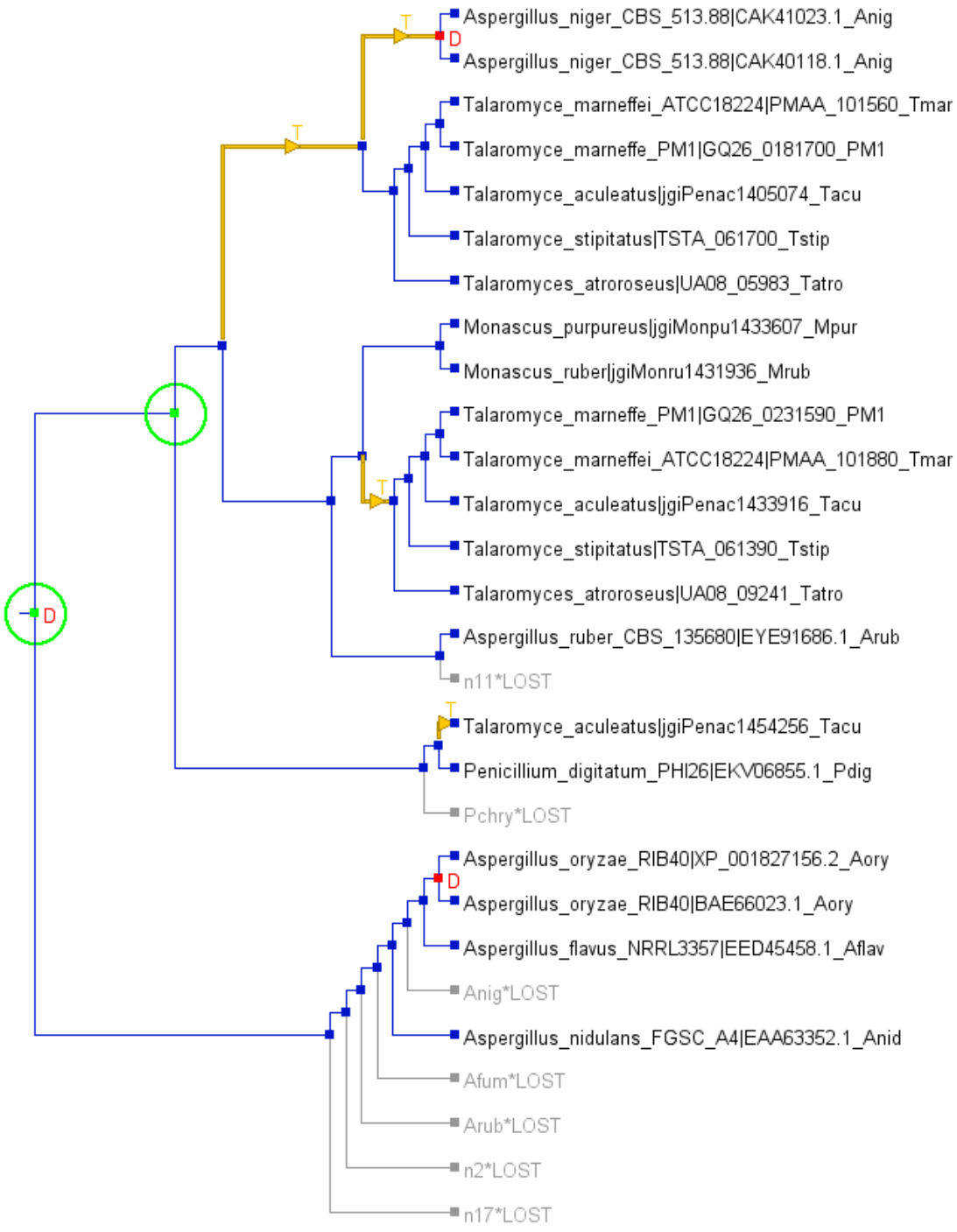


Figure D.7. Reconciled UA08_05983 gene tree #6 using Notung v2.8. Maximum parsimonious reconciliation tree #6 for UA08_05983 using duplication cost of 2, transfer cost of 4 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. These solutions are shown in UA08_05983 gene trees #5,7 shown in Figure F6,F8.

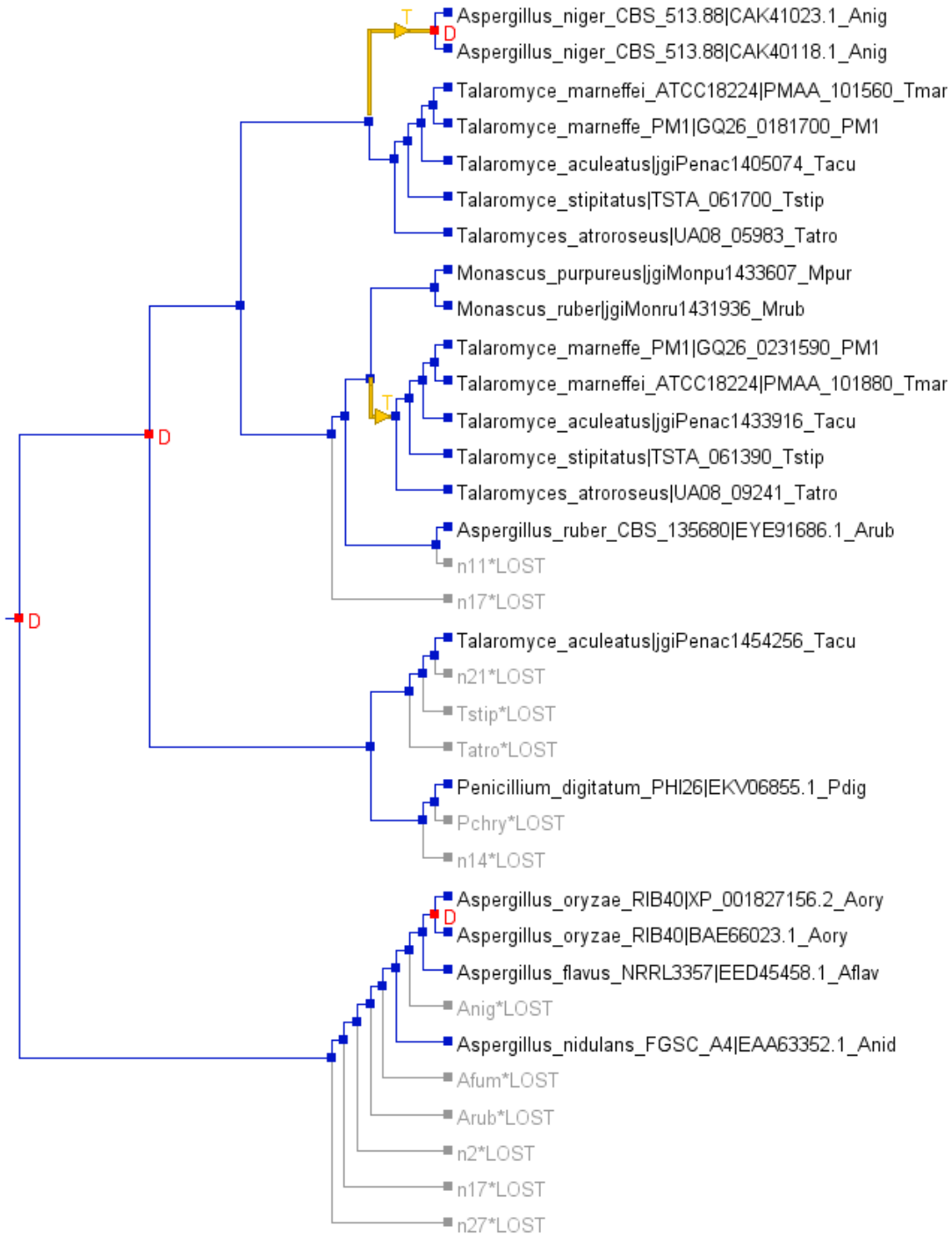


Figure D.8. Reconciled UA08_05983 gene tree #7 using Notung v2.8. Maximum parsimonious reconciliation tree #7 for UA08_05983 using the following combinations of event costs; (D,T,L) = (2,4,1), (2,5,1), (2,6,1), (4,6,1), (4,8,1), (4,10,1). Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches.

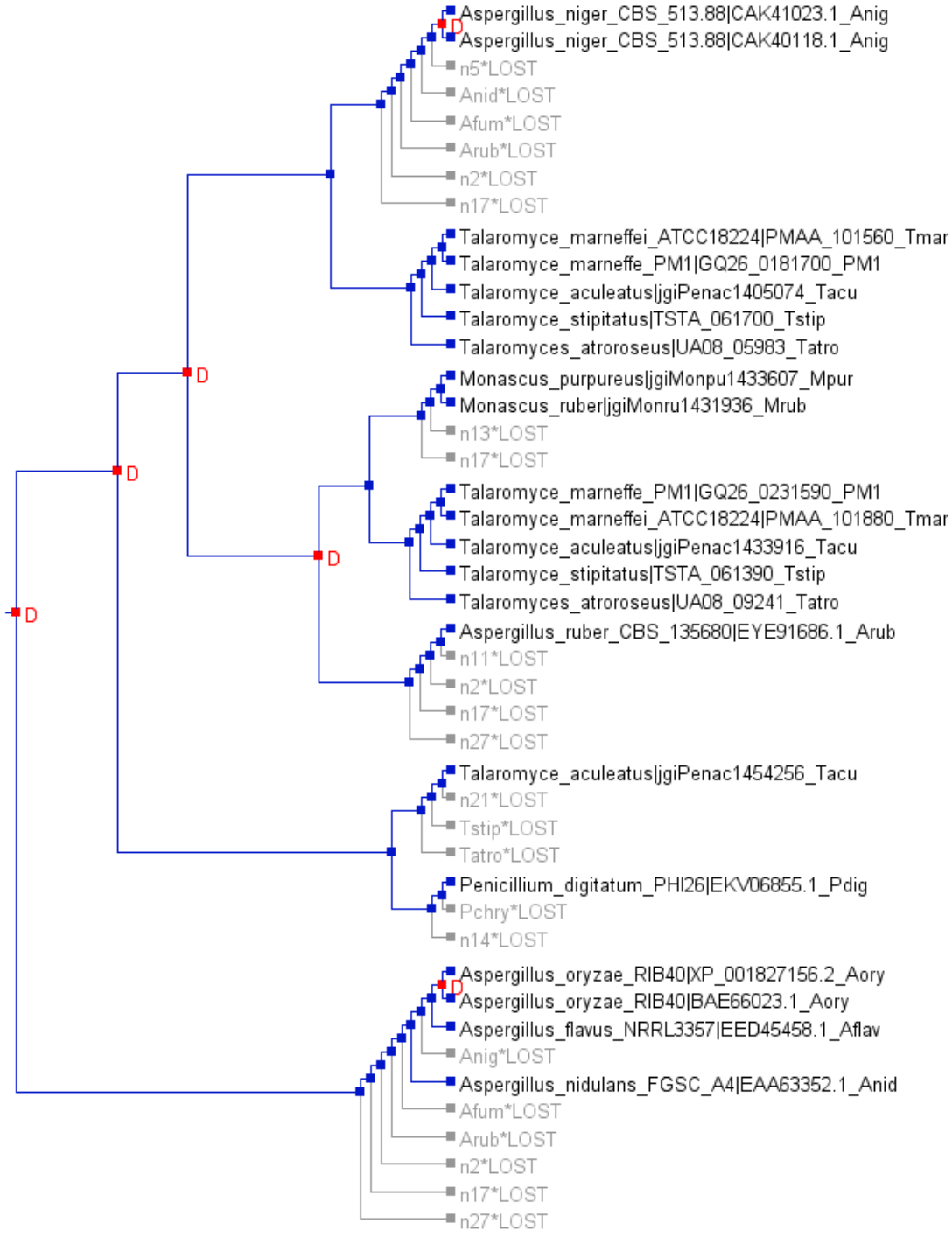


Figure D.9. Reconciled UA08_05983 gene tree #8 using Notung v2.8. Maximum parsimonious reconciliation tree #8 for UA08_05983 using the following combinations of event costs; (D,T,L) = (2,7,1), (2,9,1), (2,11,1), (4,10,1), (4,14,1), (6,14,1). Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches.

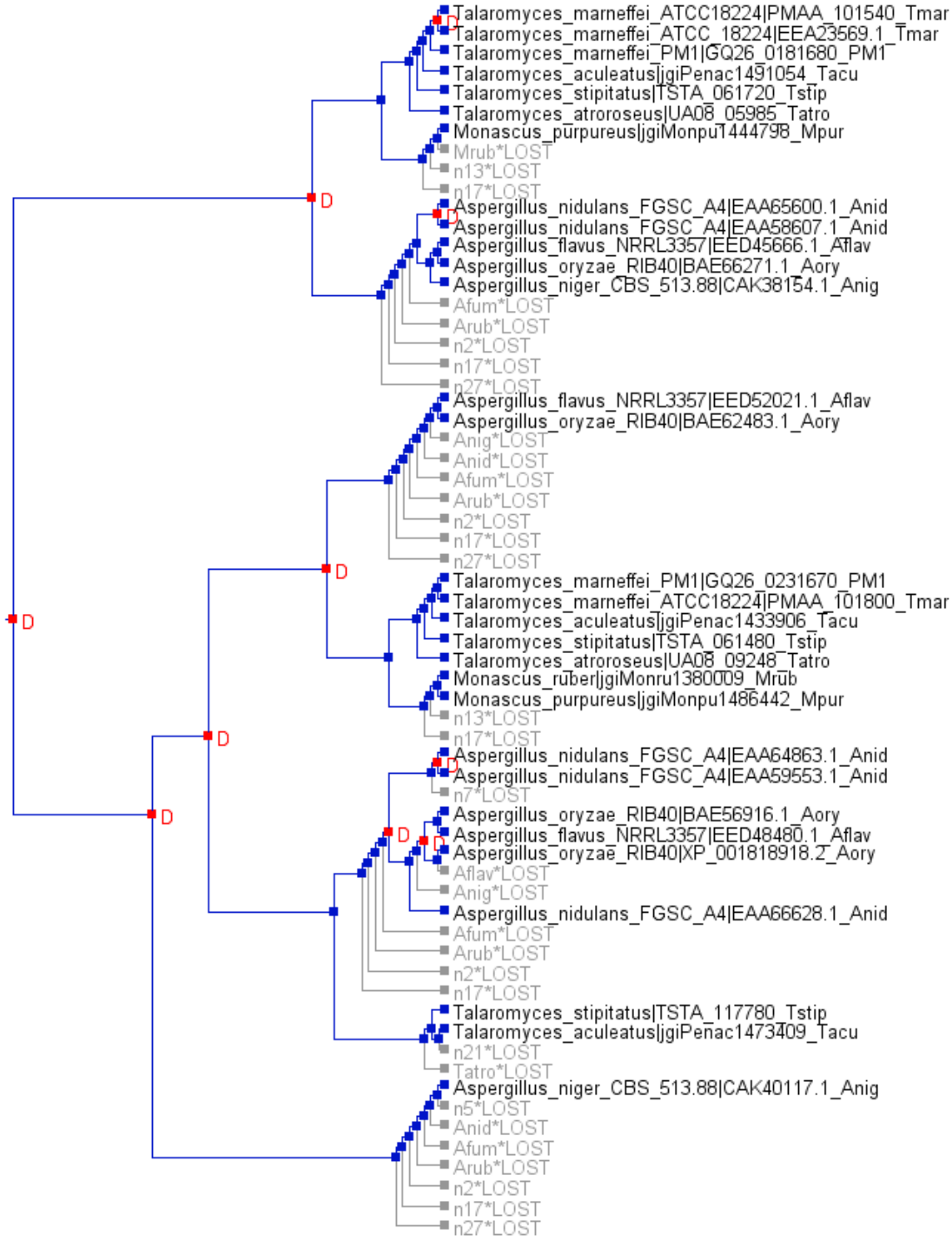


Figure D10. Reconciled UA08_05985 gene tree #1 using Notung v2.8. All combinations of DTL costs applied that produced temporal feasible solutions ended with the same reconciliation. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. The (4,8,1) event cost combination for (D,T,L)-events produced five additional equal parsimonious solutions. These solutions are shown in UA08_05985 gene trees #2-6 shown in Figure F11-15.

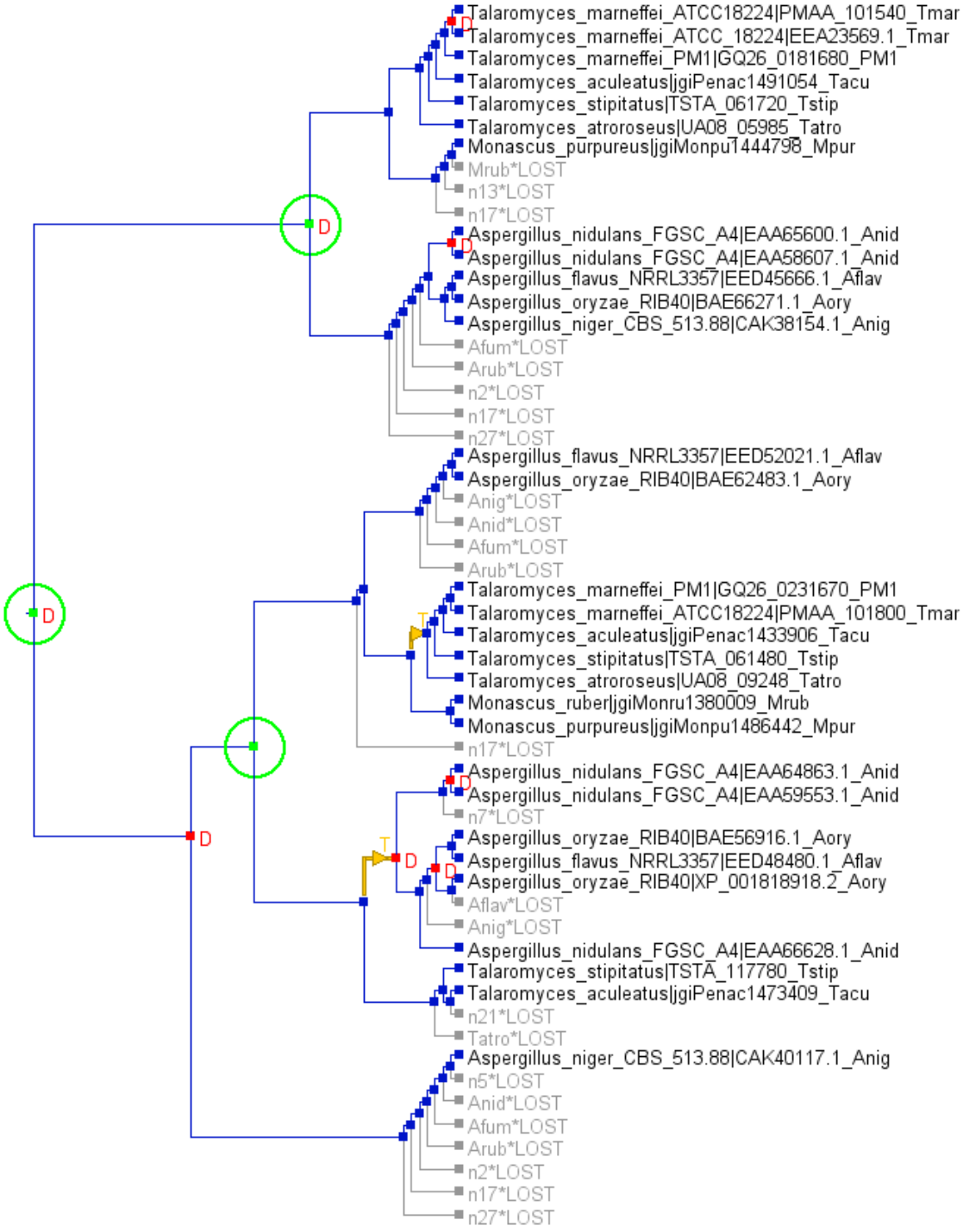


Figure D.11. Reconciled UA08_05985 gene tree #2 using Notung v2.8. Maximum parsimonious reconciliation tree #2 for UA08_05985 using duplication cost of 4, transfer cost of 8 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. These solutions are shown in UA08_05985 gene trees #1,3-6 shown in Figure F10,F12-15.

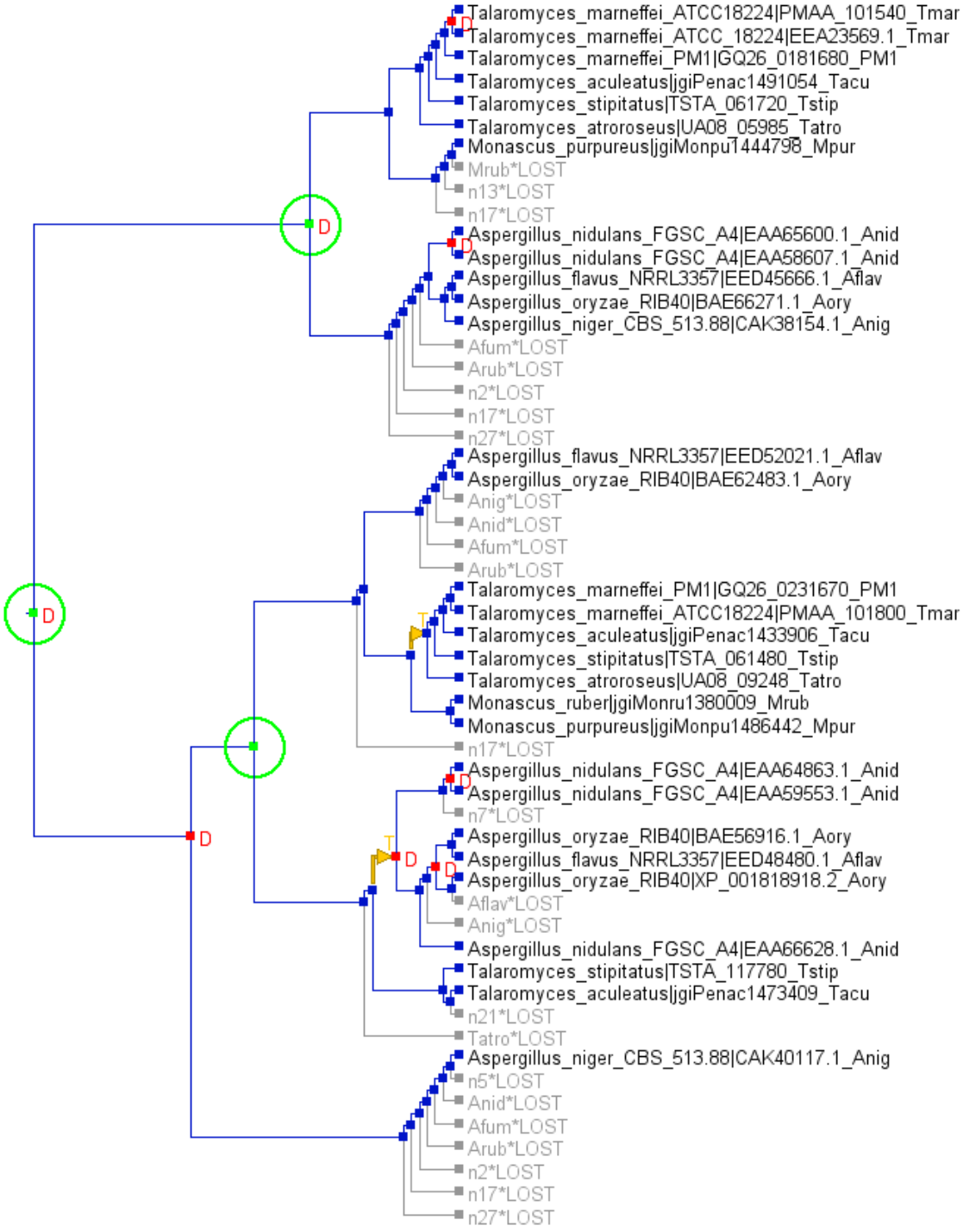


Figure D.12. Reconciled UA08_05985 gene tree #3 using Notung v2.8. Maximum parsimonious reconciliation tree #3 for UA08_05985 using duplication cost of 4, transfer cost of 8 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. These solutions are shown in UA08_05985 gene trees #1-2,4-6 shown in Figure F10-11,F13-15.

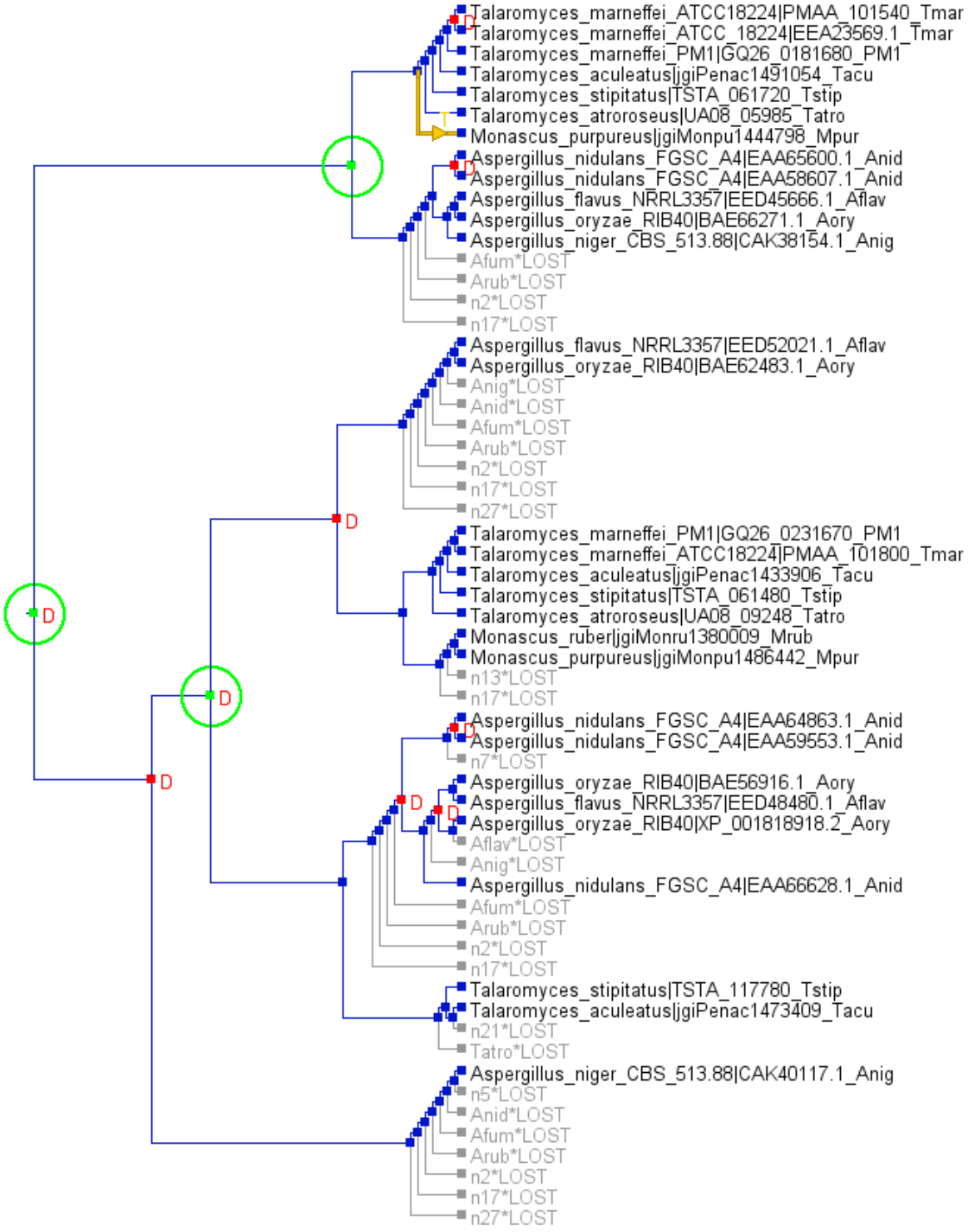


Figure D.13. Reconciled UA08_05985 gene tree #4 using Notung v2.8. Maximum parsimonious reconciliation tree #4 for UA08_05985 using duplication cost of 4, transfer cost of 8 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. These solutions are shown in UA08_05985 gene trees #1-3,5-6 shown in Figure F10-F12,F14-15.

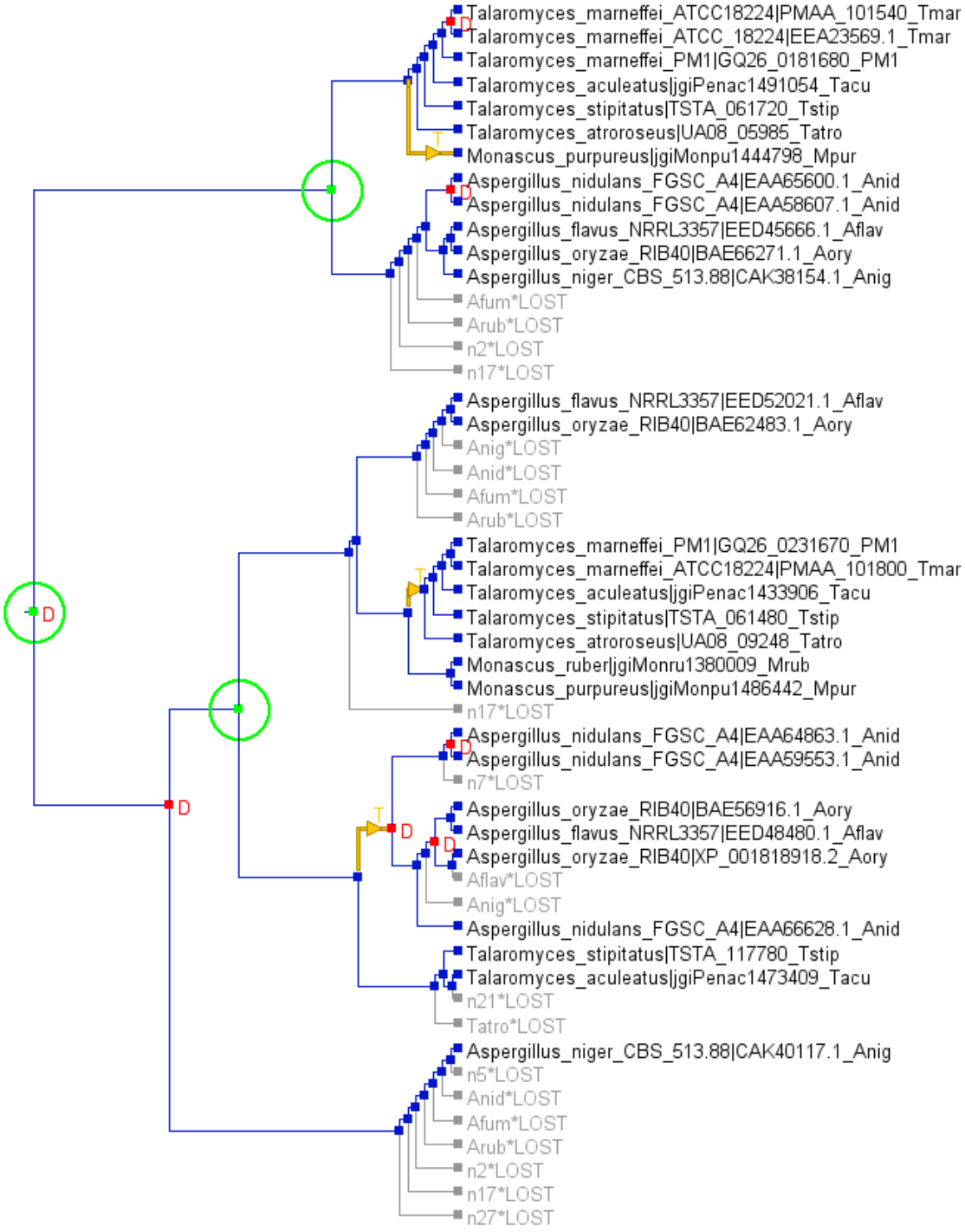


Figure D.14. Reconciled UA08_05985 gene tree #5 using Notung v2.8. Maximum parsimonious reconciliation tree #5 for UA08_05985 using duplication cost of 4, transfer cost of 8 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. These solutions are shown in UA08_05985 gene trees #1-4,6 shown in Figure F10-F13,15.

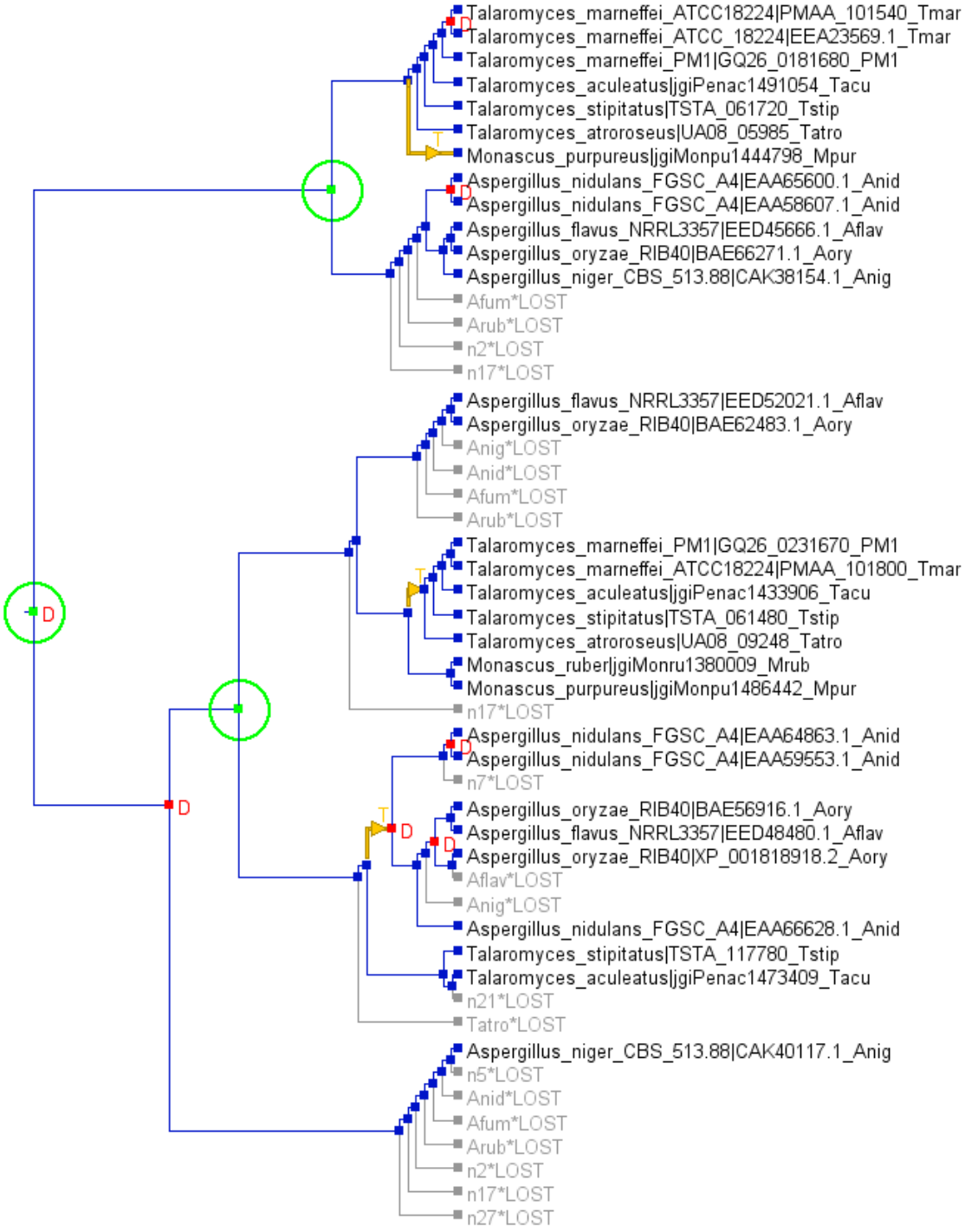


Figure D.15. Reconciled UA08_05985 gene tree #6 using Notung v2.8. Maximum parsimonious reconciliation tree #6 for UA08_05985 using duplication cost of 4, transfer cost of 8 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. These solutions are shown in UA08_05985 gene trees #1-5 shown in Figure F10-F14.

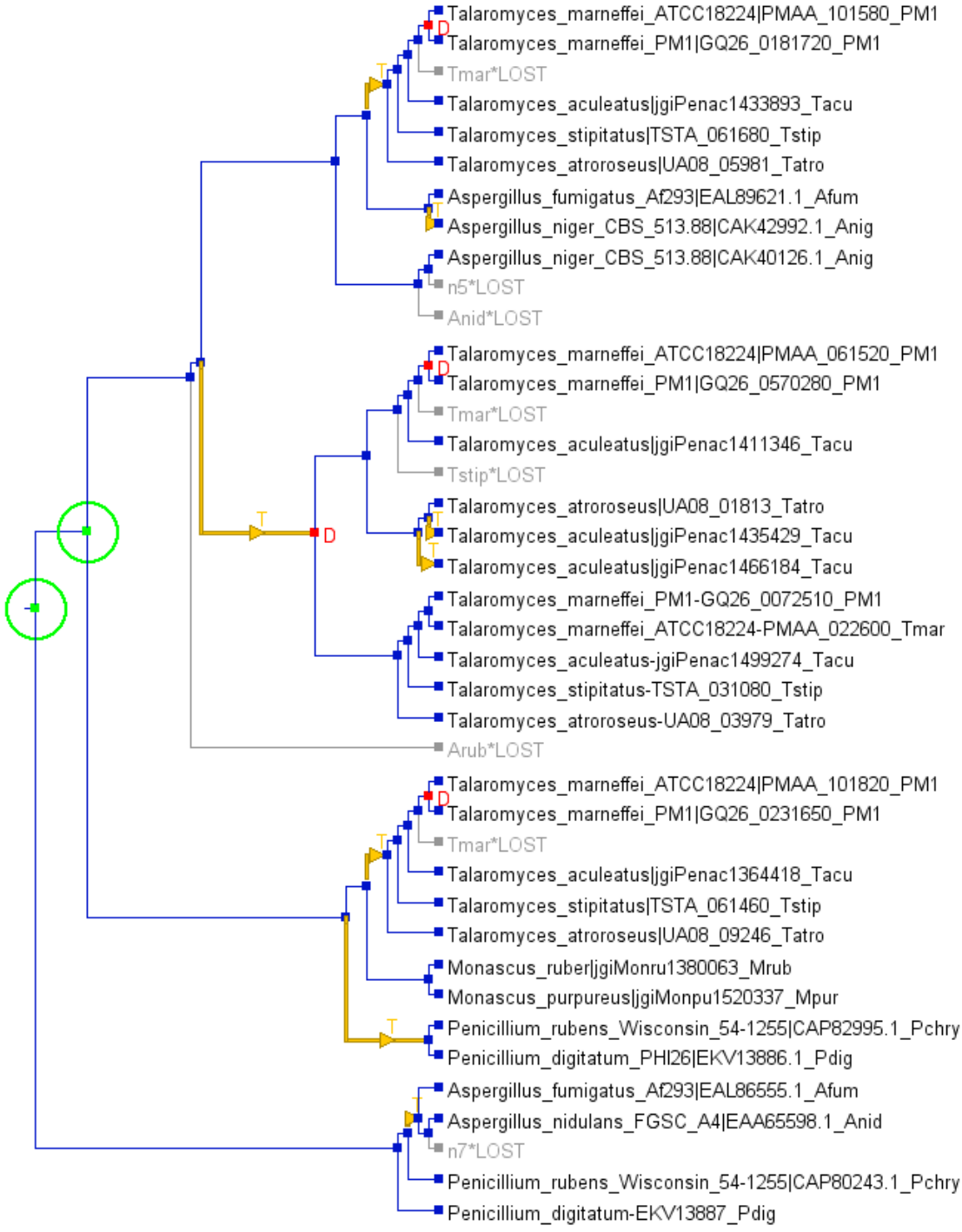


Figure D.16. Reconciled UA08_05981 gene tree #1 using Notung v2.8. Maximum parsimonious reconciliation tree #1 for UA08_05981 using duplication cost of 2, transfer cost of 3 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. The additional solution are shown in UA08_05981 gene tree #2 shown in Figure F17.

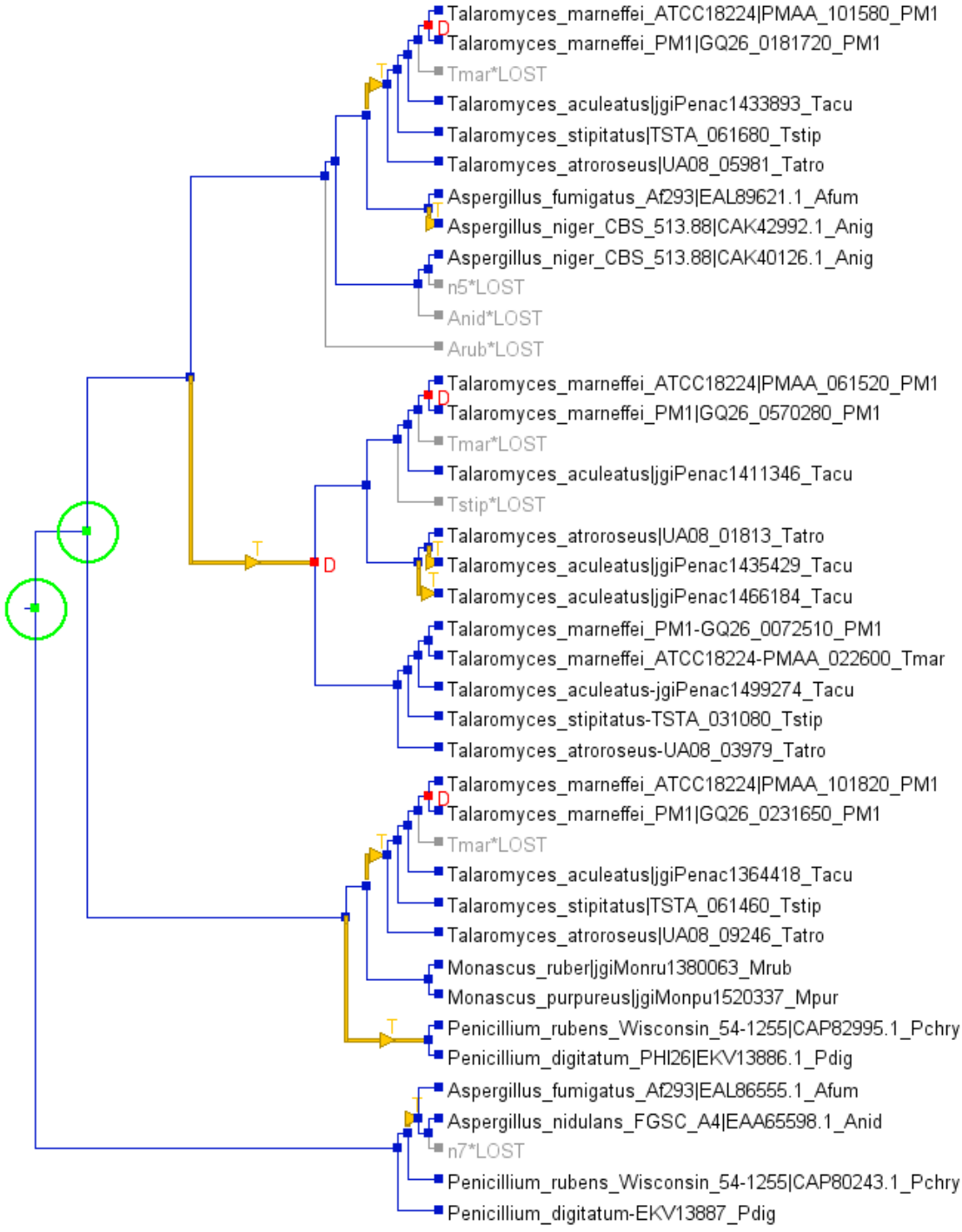


Figure D.17. Reconciled UA08_05981 gene tree #2 using Notung v2.8. Maximum parsimonious reconciliation tree #2 for UA08_05981 using duplication cost of 2, transfer cost of 3 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. The additional solution is shown in UA08_05981 gene tree #1 shown in Figure F16.

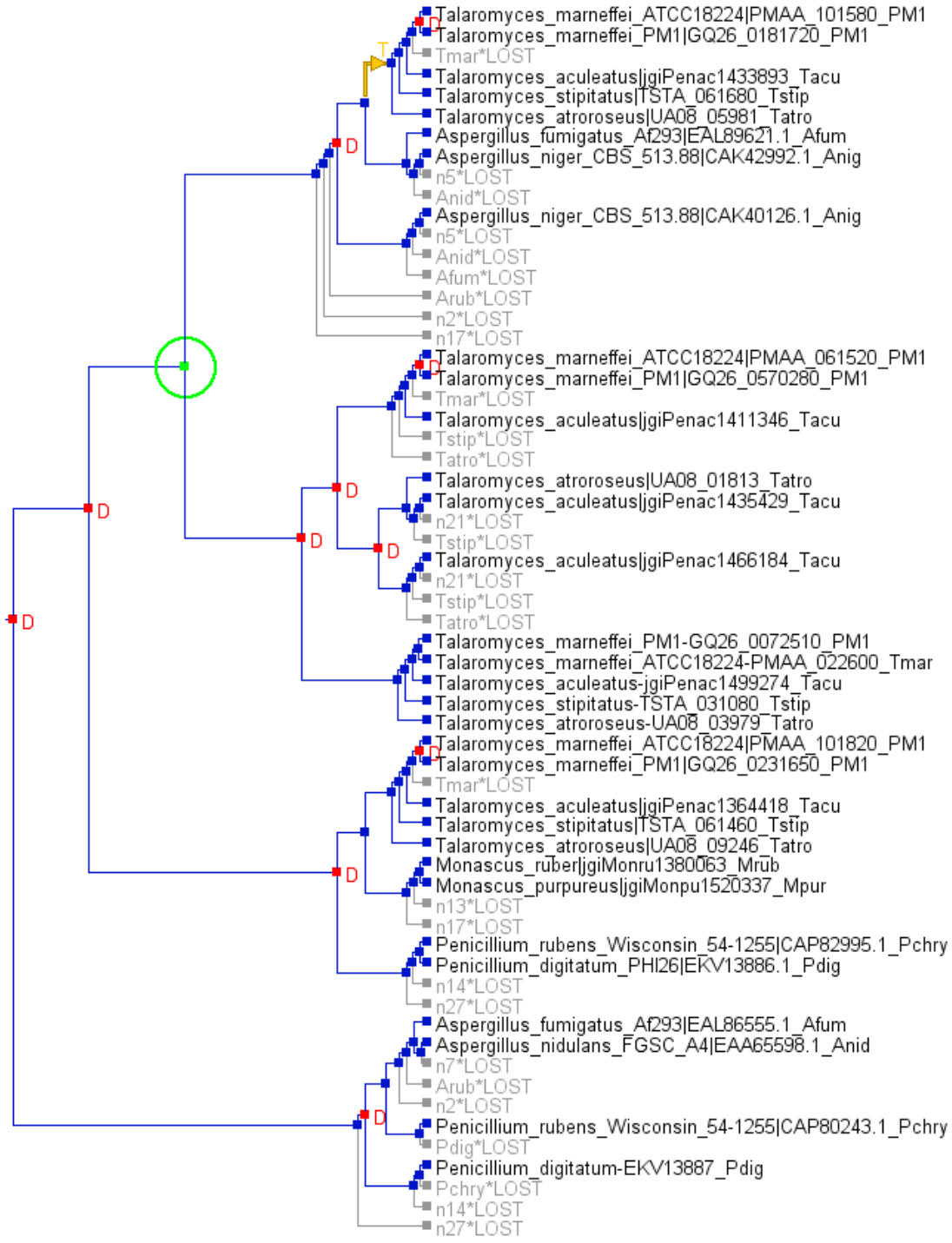


Figure D.18. Reconciled UA08_05981 gene tree #3 using Notung v2.8. Maximum parsimonious reconciliation tree #3 for UA08_05981 using duplication cost of 2, transfer cost of 7 and loss cost of 1. Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches. Green circles display nodes where multiple equal parsimonious solutions to the subtree exists. The additional solution is shown in UA08_05981 gene tree #4 shown in Figure F19.

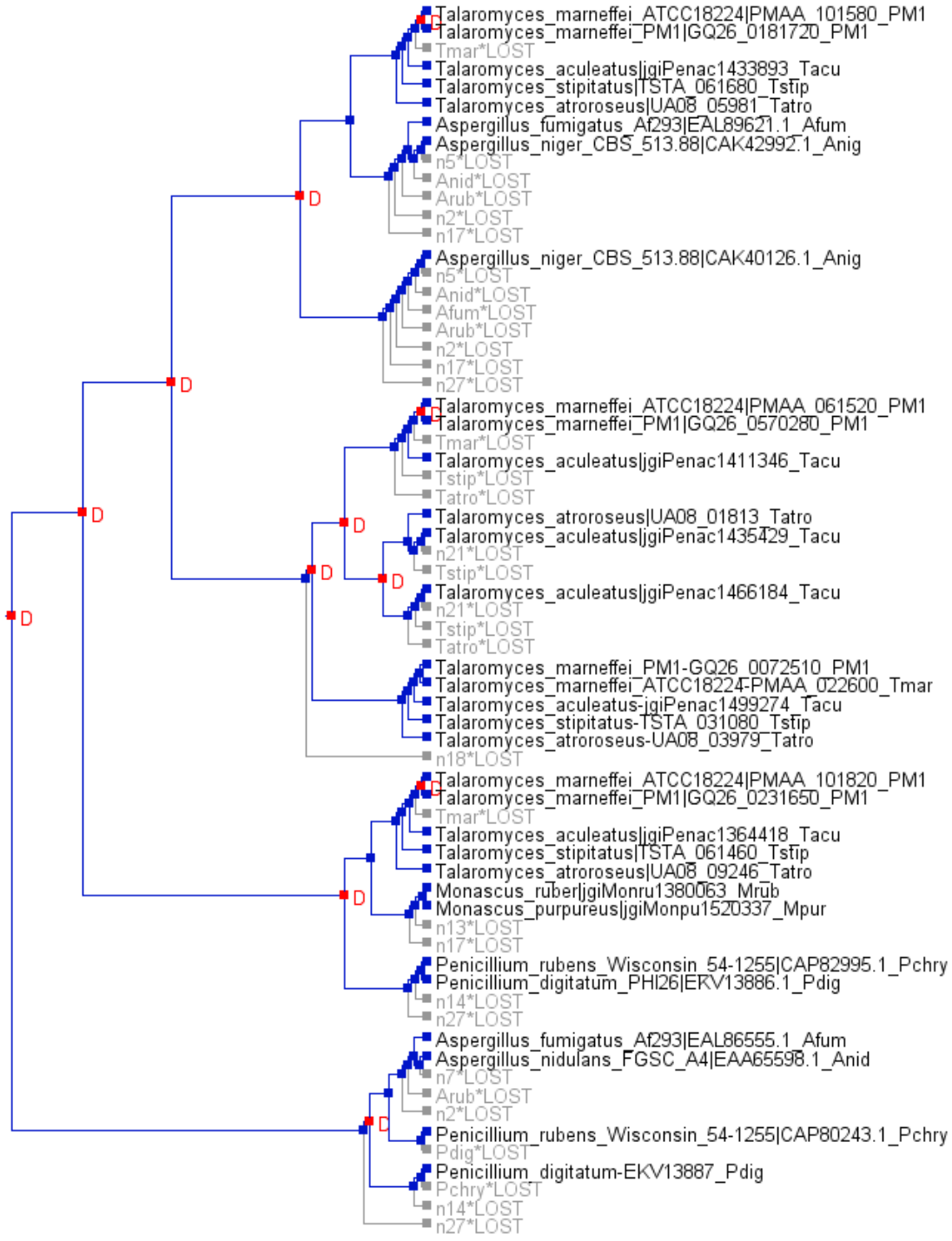


Figure D.19. Reconciled UA08_05981 gene tree #4 using Notung v2.8. Maximum parsimonious reconciliation tree #4 for UA08_05981 using the following combinations of event costs; (D,T,L) = (2,7,1), (2,9,1), (2,11,1), (4,10,1), (4,14,1), (6,14,1). Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches.

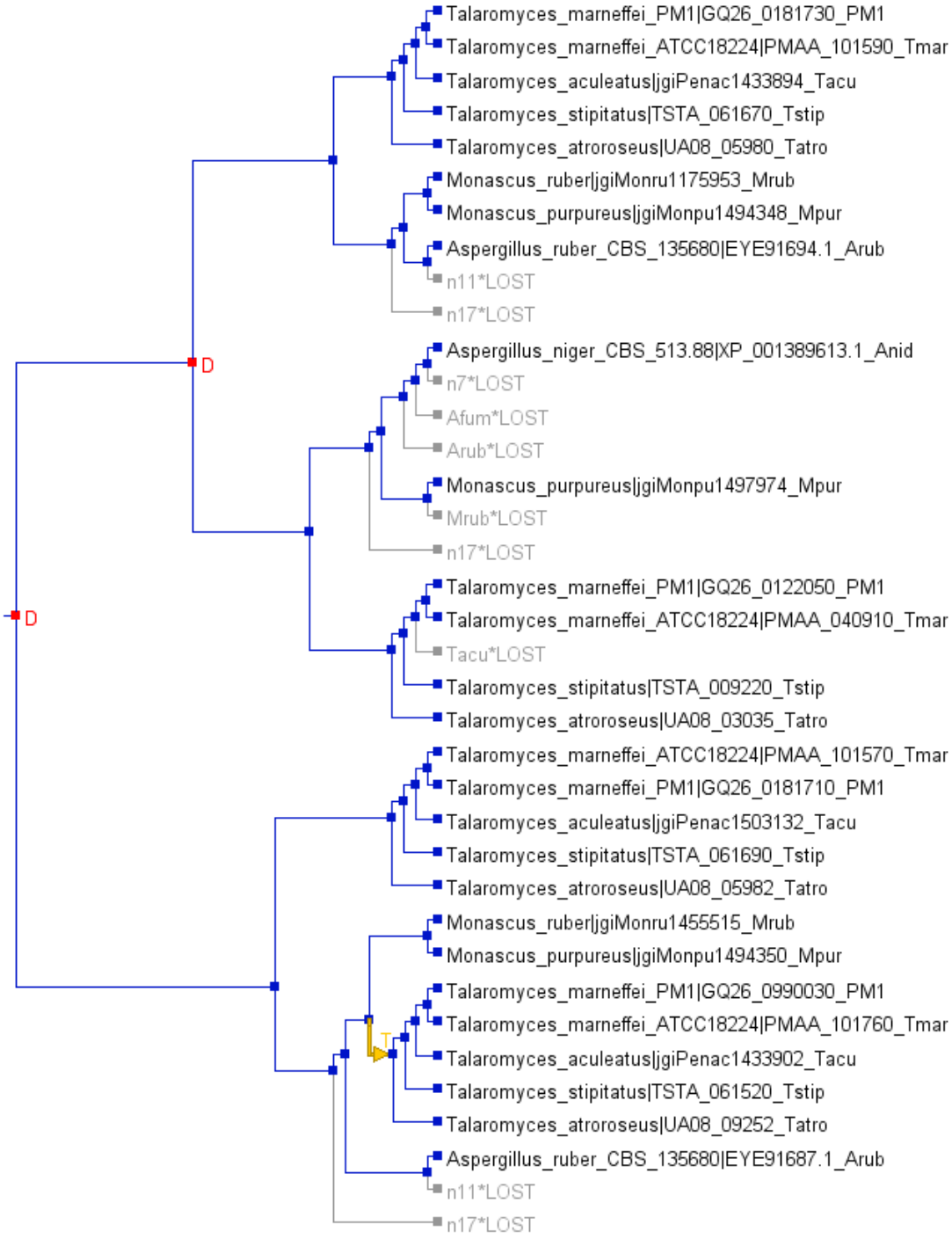


Figure D.20. Reconciled UA08_05982 gene tree #1 using Notung v2.8. Maximum parsimonious reconciliation tree #1 for UA08_05982 using the following combinations of event costs; (D,T,L) = (2,3,1), (2,4,1), (2,5,1), (2,7,1), (4,6,1), (4,10,1), (6,10,1), (6,14,1). Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches.

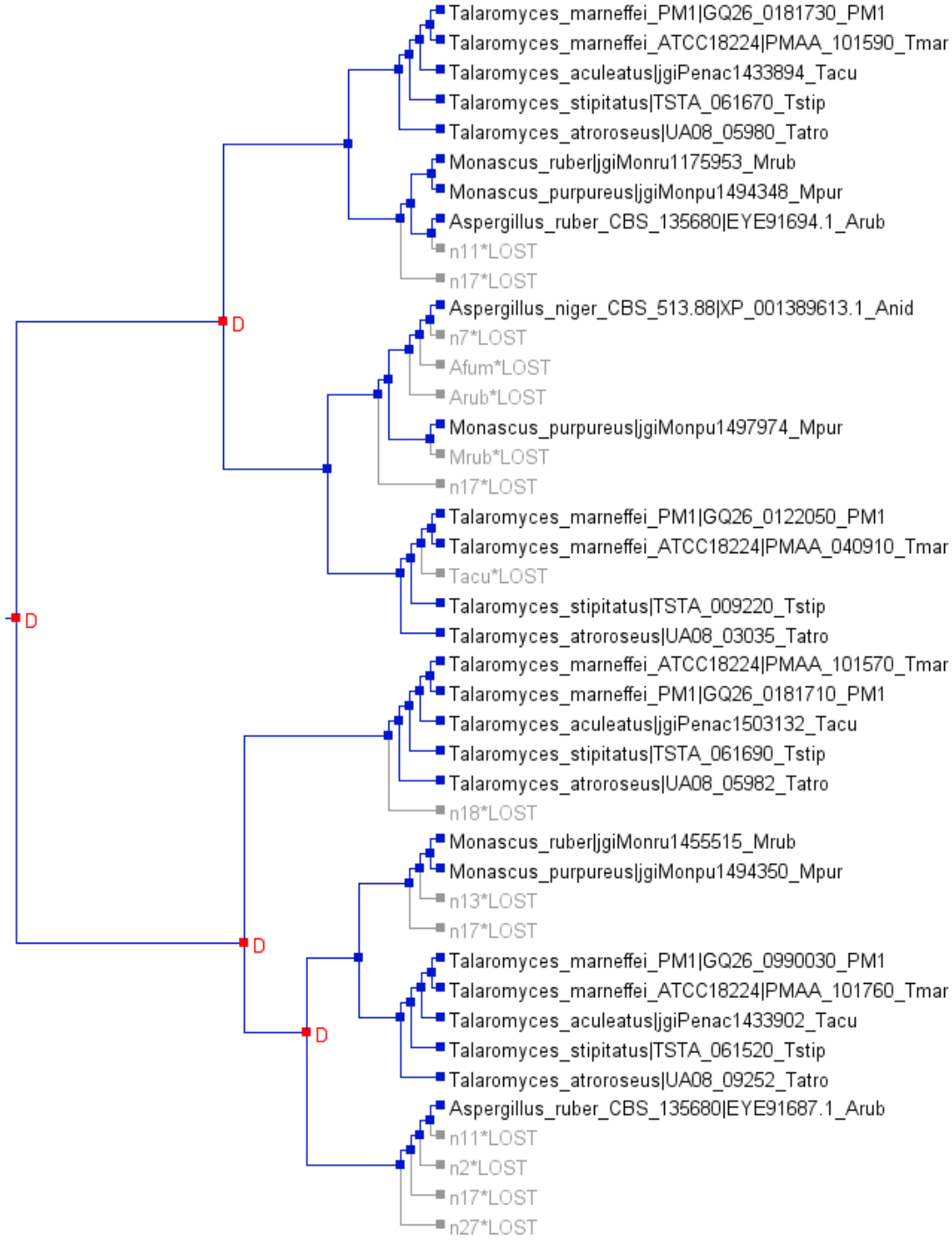


Figure D.21. Reconciled UA08_05982 gene tree #2 using Notung v2.8. Maximum parsimonious reconciliation tree #2 for UA08_05982 using the following combinations of event costs; (D,T,L) = (2,9,1), (2,11,1), (4,14,1), (6,10,1), (6,14,1). Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches. Transfer events are shown as yellow branches.

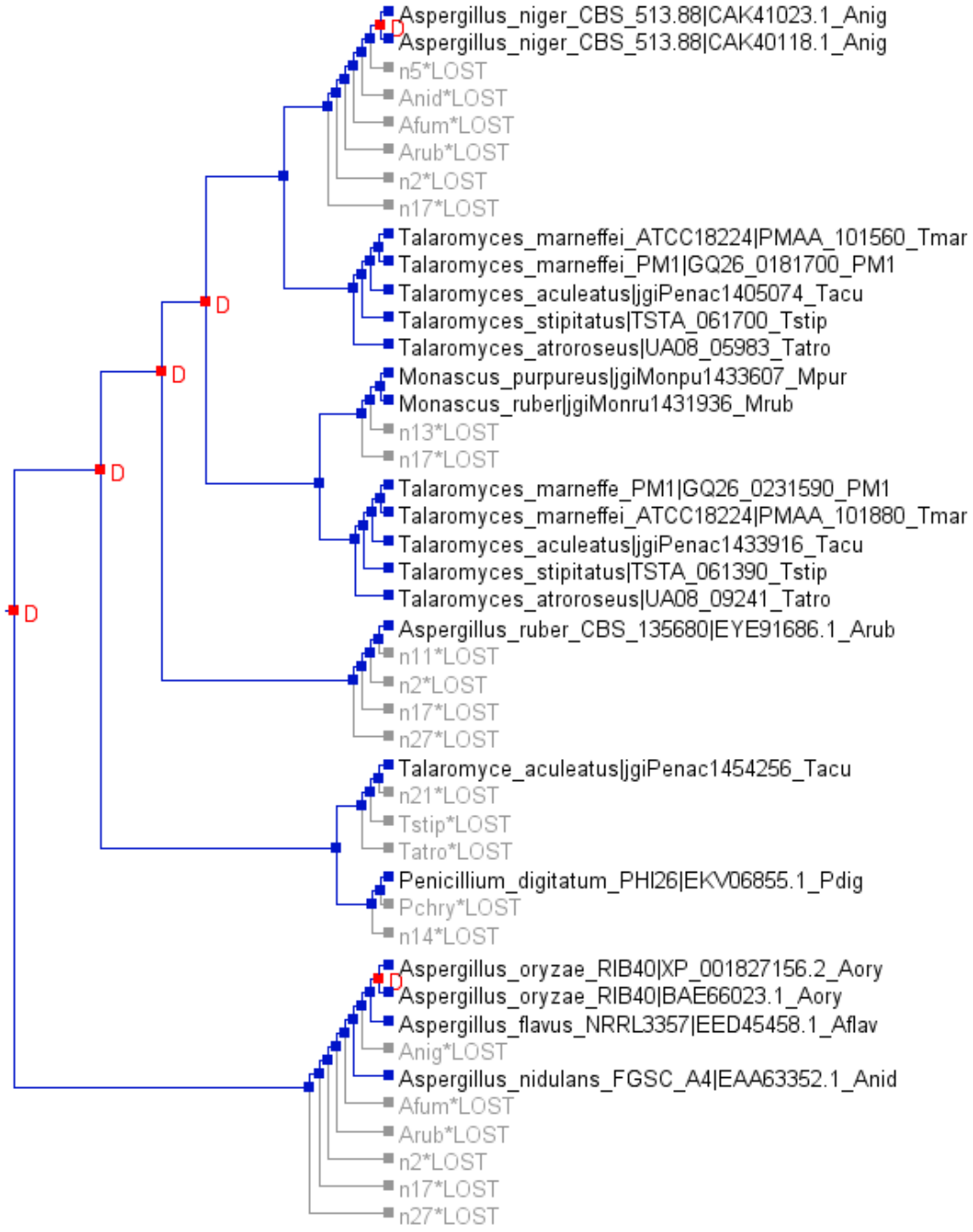


Figure D.22. Reconciled UA08_05983 gene tree with moved *A. ruber* branch. When the low supported branch of *A. ruber* is moved to before the node were the mitorubrin and MP clades branch the reconciliation of the following combinations of event costs change to favour a strictly duplication-loss evolutionary scenario; (D,T,L) = (2,7,1), (4,8,1), (6,14,1). Duplication events are highlighted at internal nodes with a red D. Gene losses are shown as light grey branches.

Appendix E

Supporting information to Chapter 7

ORF1

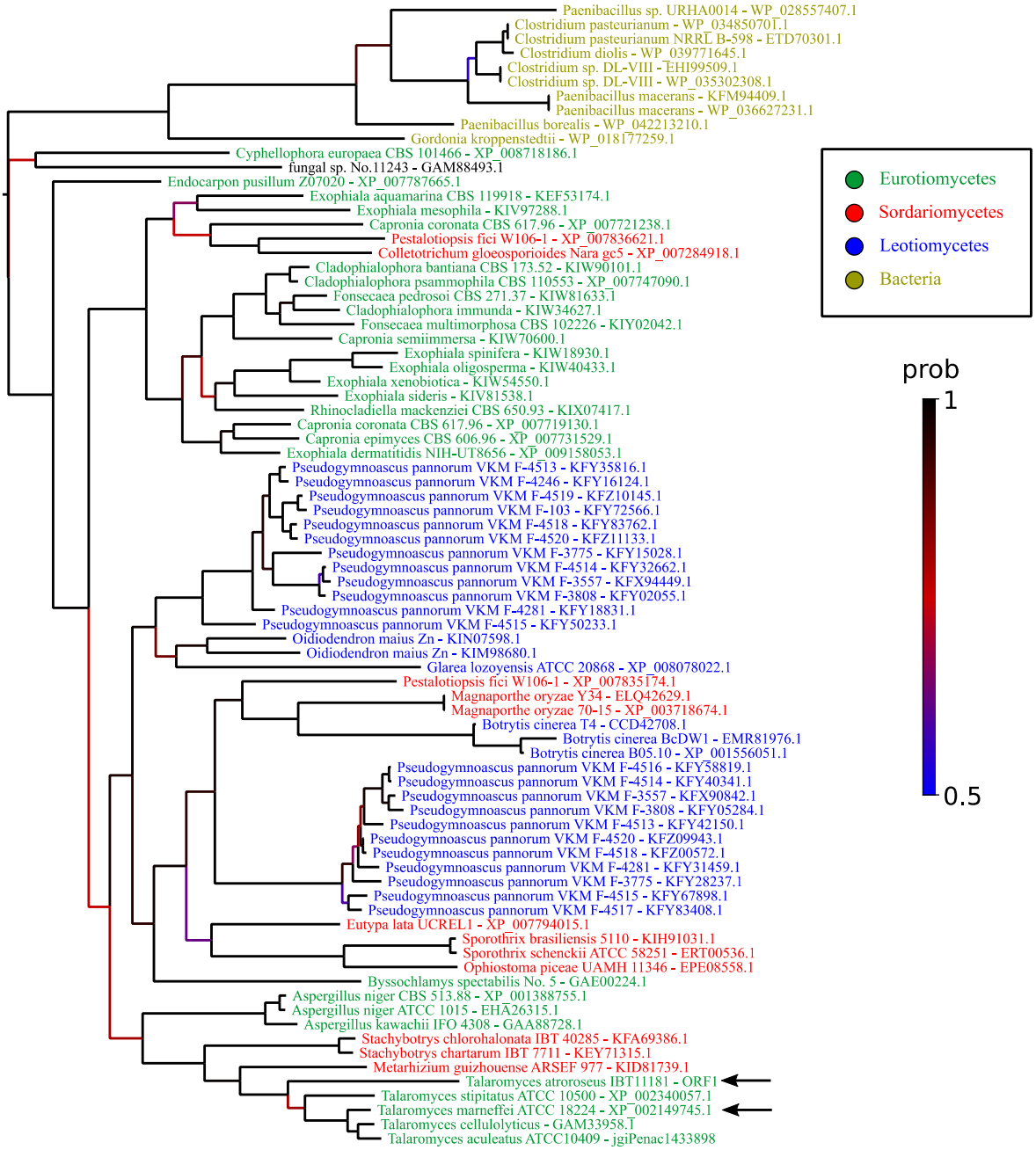


Figure E.1. Phylogenetic relationship of *T.* atroroseus IBT11181 mitorubrin ORF1 homologs. The tree is a 50% majority rule consensus constructed with MrBayes. Branches are coloured according to its posterior probability, see colour gradient scale. Arrows point toward proteins coming from genes appearing in the azaphilone clusters in figure 7.1. Each protein has been coloured according to the taxonomical class of the species it is coming from (see figure legend).

ORF4

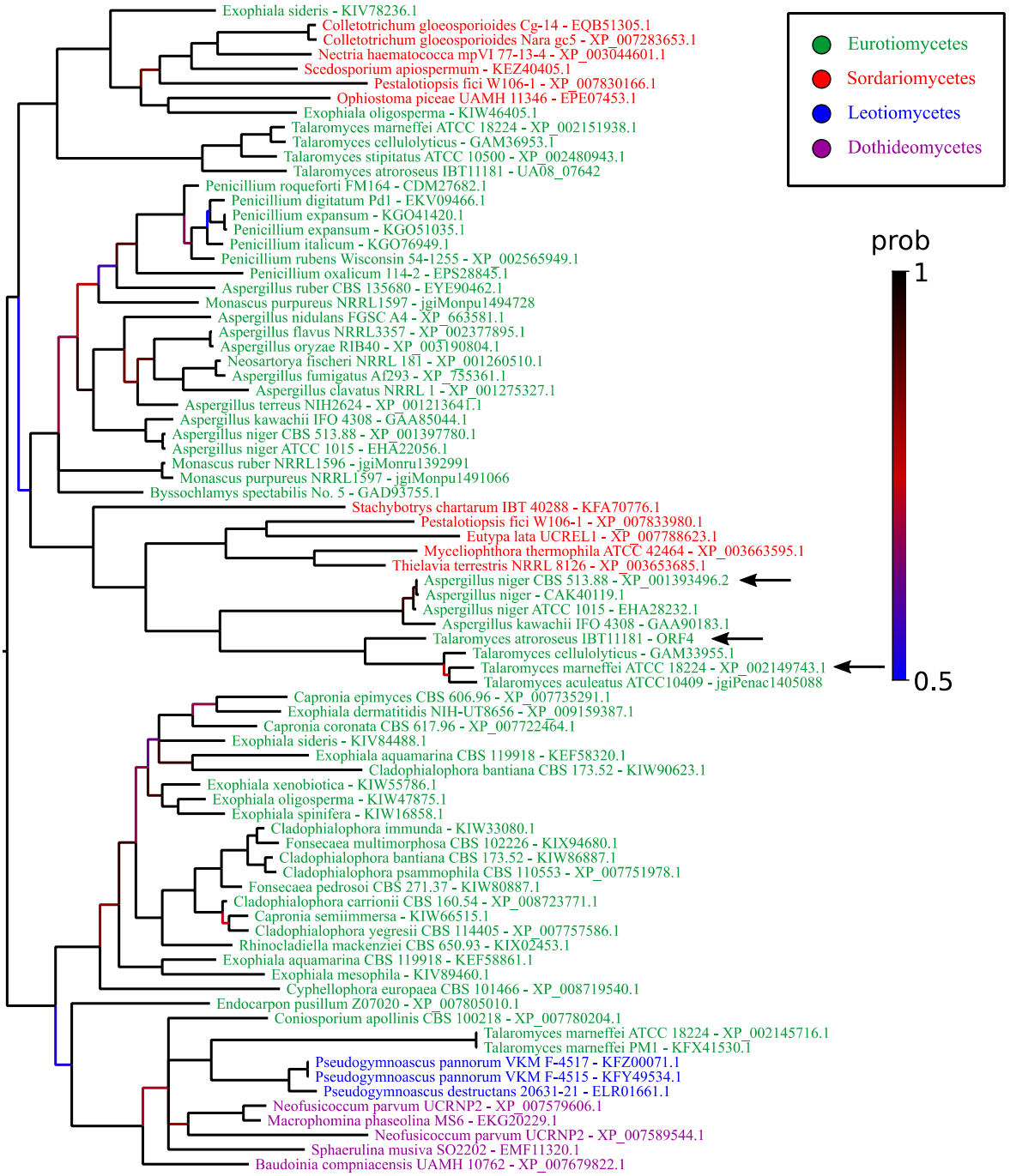


Figure E.2. Phylogenetic relationship of *T.* atroroseus IBT11181 mitorubrin ORF4 homologs. The tree is a 50% majority rule consensus constructed with MrBayes. Branches are coloured according to its posterior probability, see colour gradient scale. Arrows point toward proteins coming from genes appearing in the azaphilone clusters in figure 7.1. Each protein has been coloured according to the taxonomical class of the species it is coming from (see figure legend).

ORF6

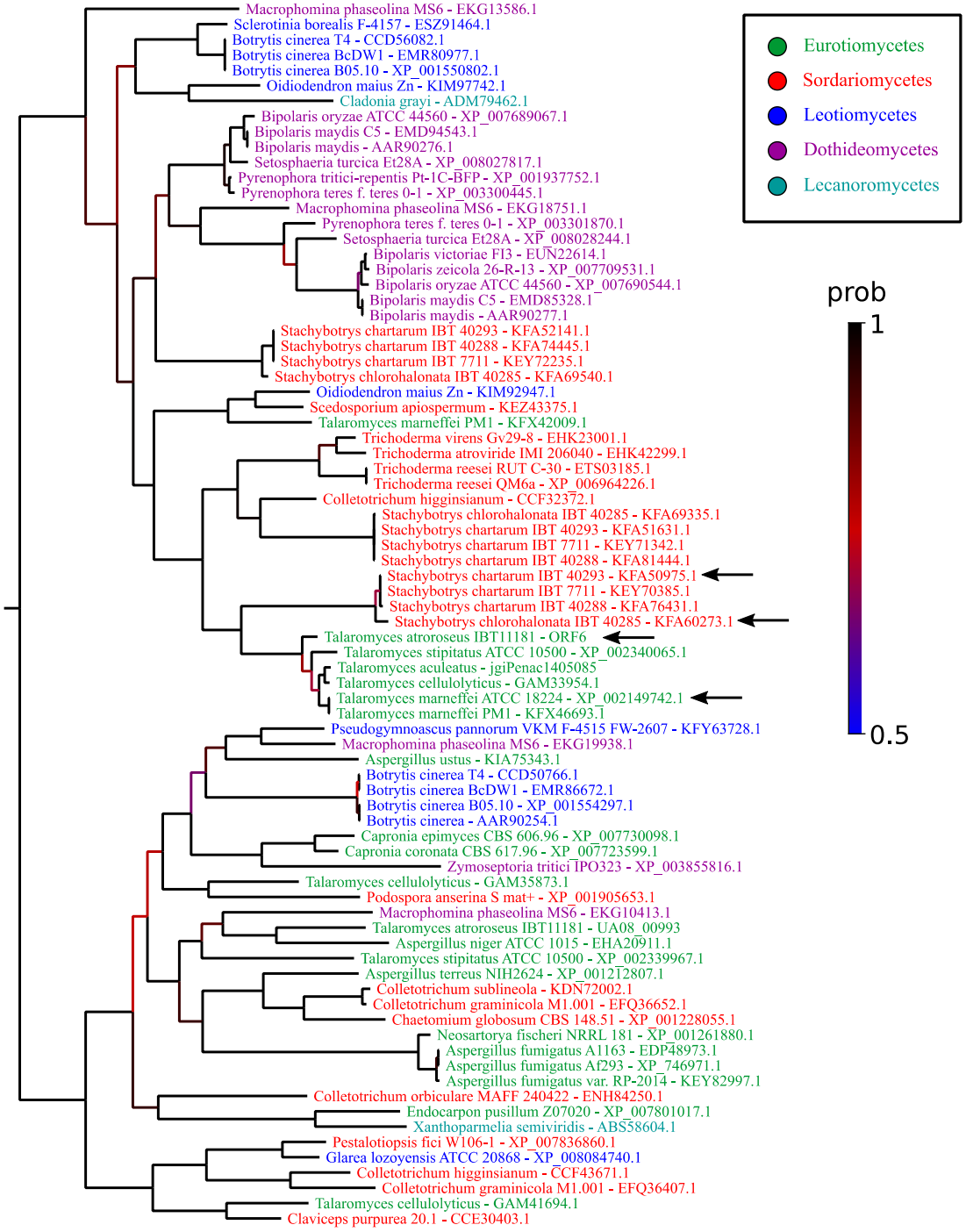


Figure E.3. Phylogenetic relationship of *T.* atroroseus IBT11181 mitorubrin ORF6 homologs. The tree is a 50% majority rule consensus constructed with MrBayes. Branches are coloured according to its posterior probability, see colour gradient scale. Arrows point toward proteins coming from genes appearing in the azaphilone clusters in figure 7.1. Each protein has been coloured according to the taxonomical class of the species it is coming from (see figure legend).

ORF7 / ORF9 tree

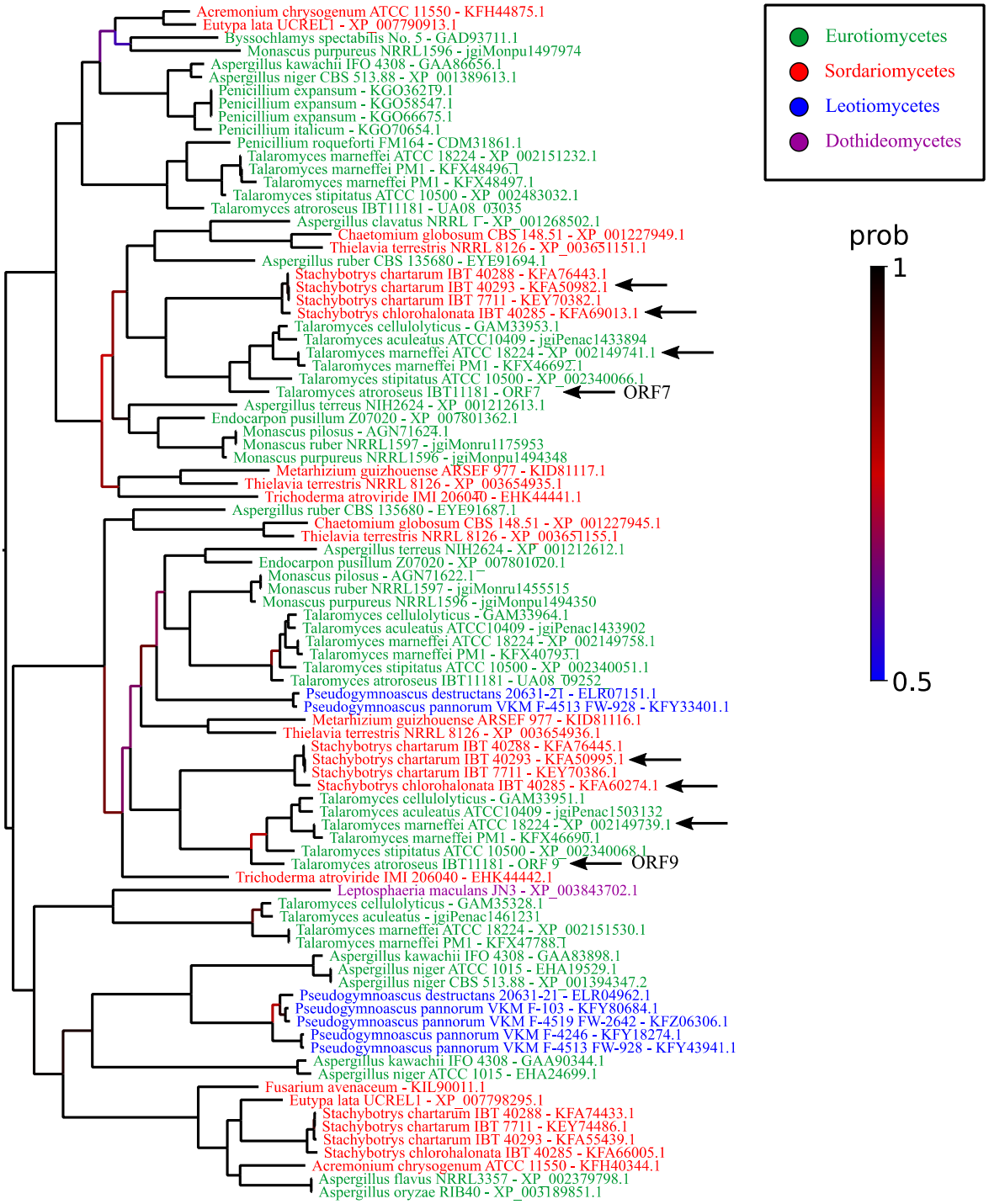


Figure E.4. Phylogenetic relationship of *T.* atroroseus IBT11181 mitorubrin ORF7 and ORF9 homologs. The tree is a 50% majority rule consensus constructed with MrBayes. Branches are coloured according to its posterior probability, see colour gradient scale. Arrows point toward proteins coming from genes appearing in the azaphilone clusters in figure 7.1. Each protein has been coloured according to the taxonomical class of the species it is coming from (see figure legend).

ORF10

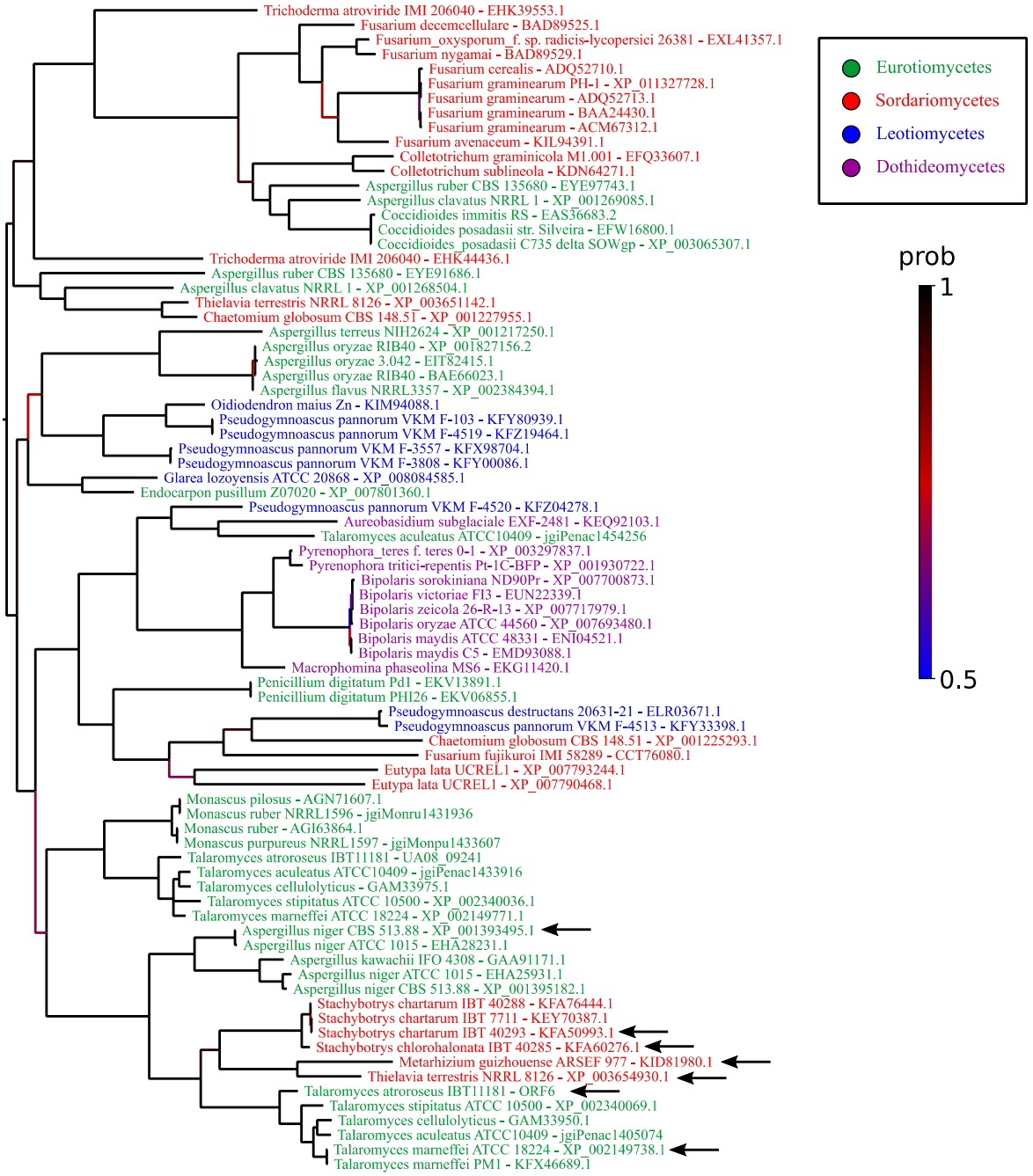


Figure E.5. Phylogenetic relationship of *T.* atroroseus IBT11181 mitorubrin ORF10 homologs. The tree is a 50% majority rule consensus constructed with MrBayes. Branches are coloured according to its posterior probability, see colour gradient scale. Arrows point toward proteins coming from genes appearing in the azaphilone clusters in figure 7.1. Each protein has been coloured according to the taxonomical class of the species it is coming from (see figure legend).

ORF11

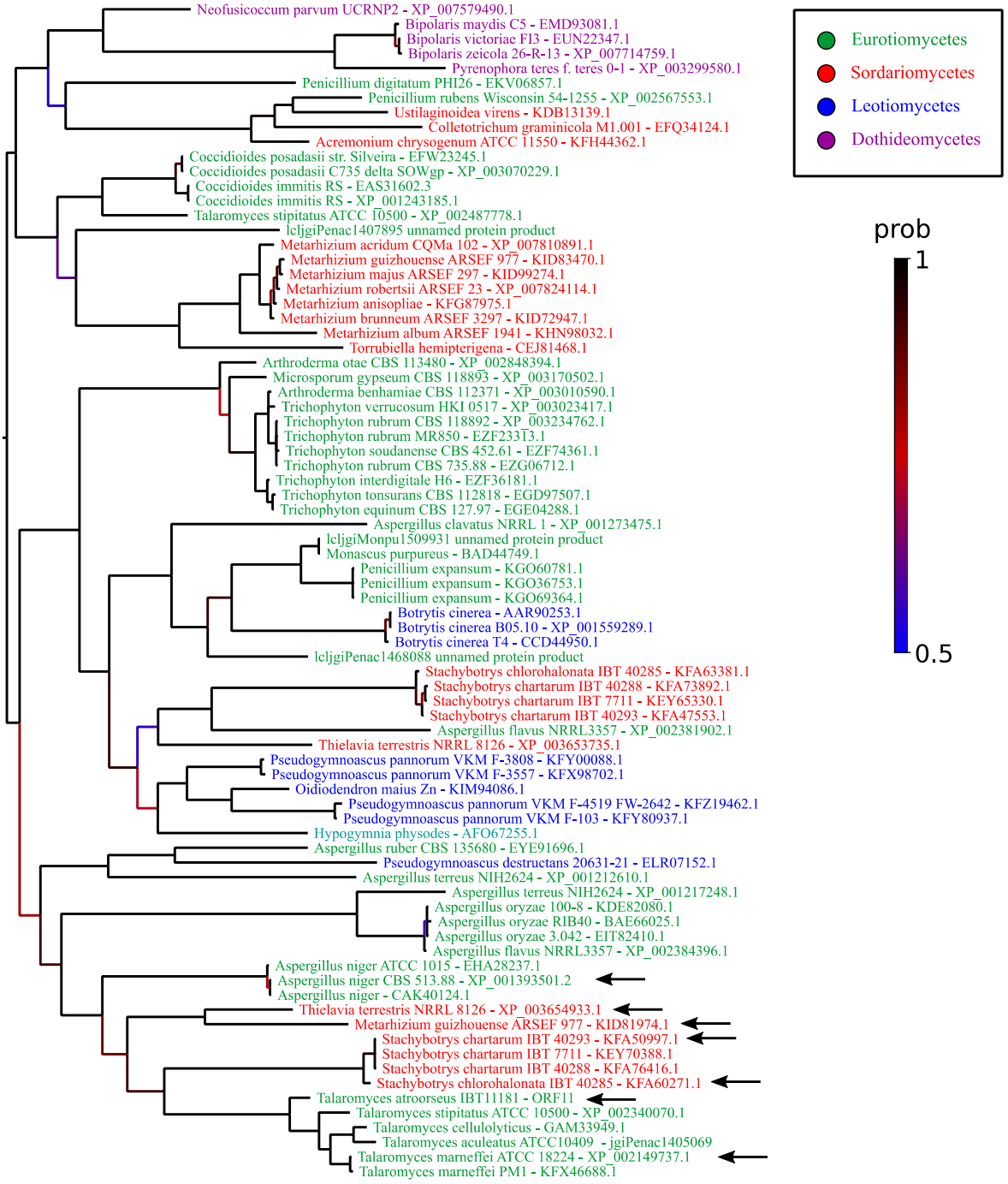


Figure E.6. Phylogenetic relationship of *T. atroroseus* IBT11181 mitorubrin ORF11 homologs. The tree is a 50% majority rule consensus constructed with MrBayes. Branches are coloured according to its posterior probability, see colour gradient scale. Arrows point toward proteins coming from genes appearing in the azaphilone clusters in figure 7.1. Each protein has been coloured according to the taxonomical class of the species it is coming from (see figure legend).

ORF12

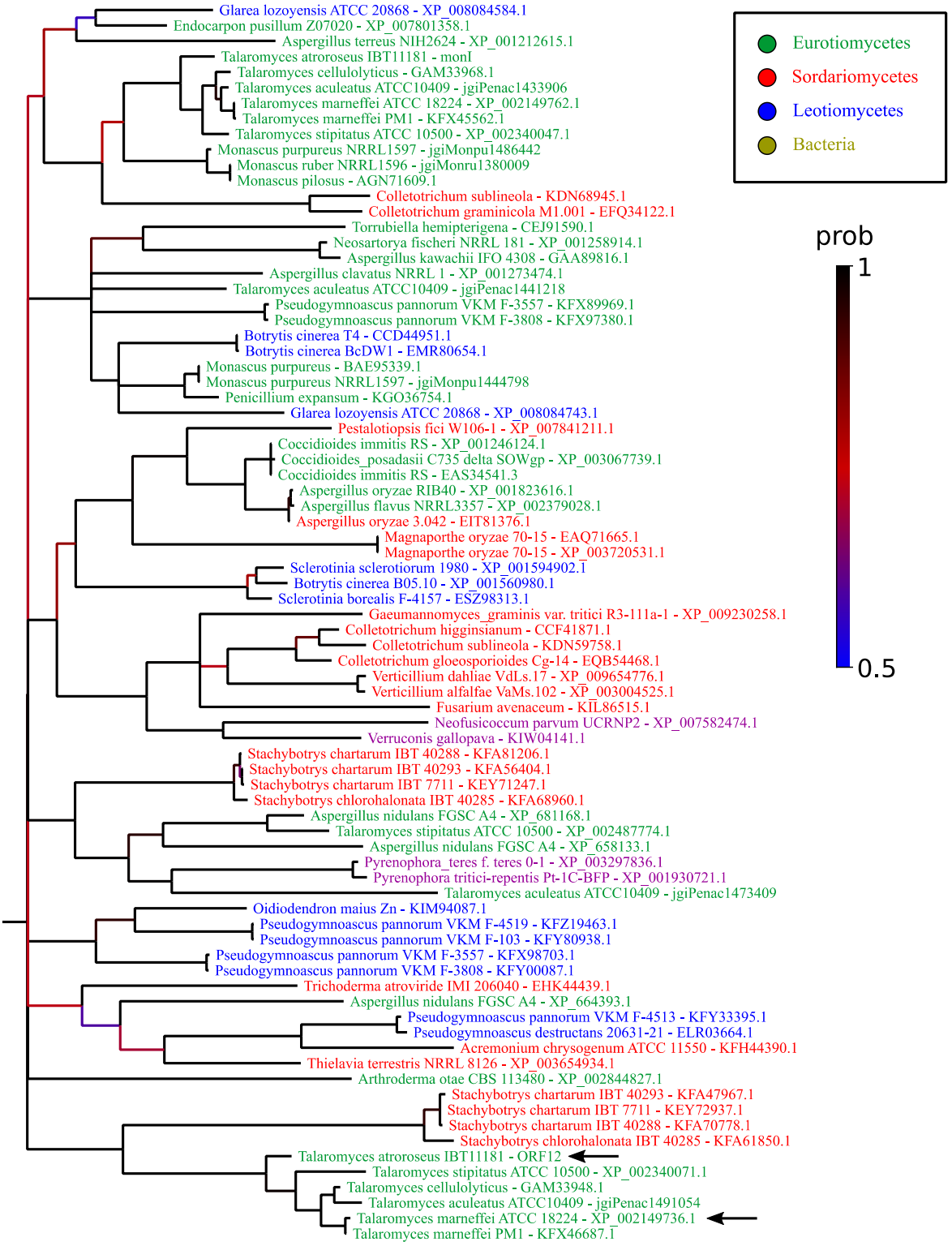


Figure E.7. Phylogenetic relationship of *T.* atroroseus IBT11181 mitorubrin ORF12 homologs. The tree is a 50% majority rule consensus constructed with MrBayes. Branches are coloured according to its posterior probability, see colour gradient scale. Arrows point toward proteins coming from genes appearing in the azaphilone clusters in figure 7.1. Each protein has been coloured according to the taxonomical class of the species it is coming from (see figure legend).

ORFB

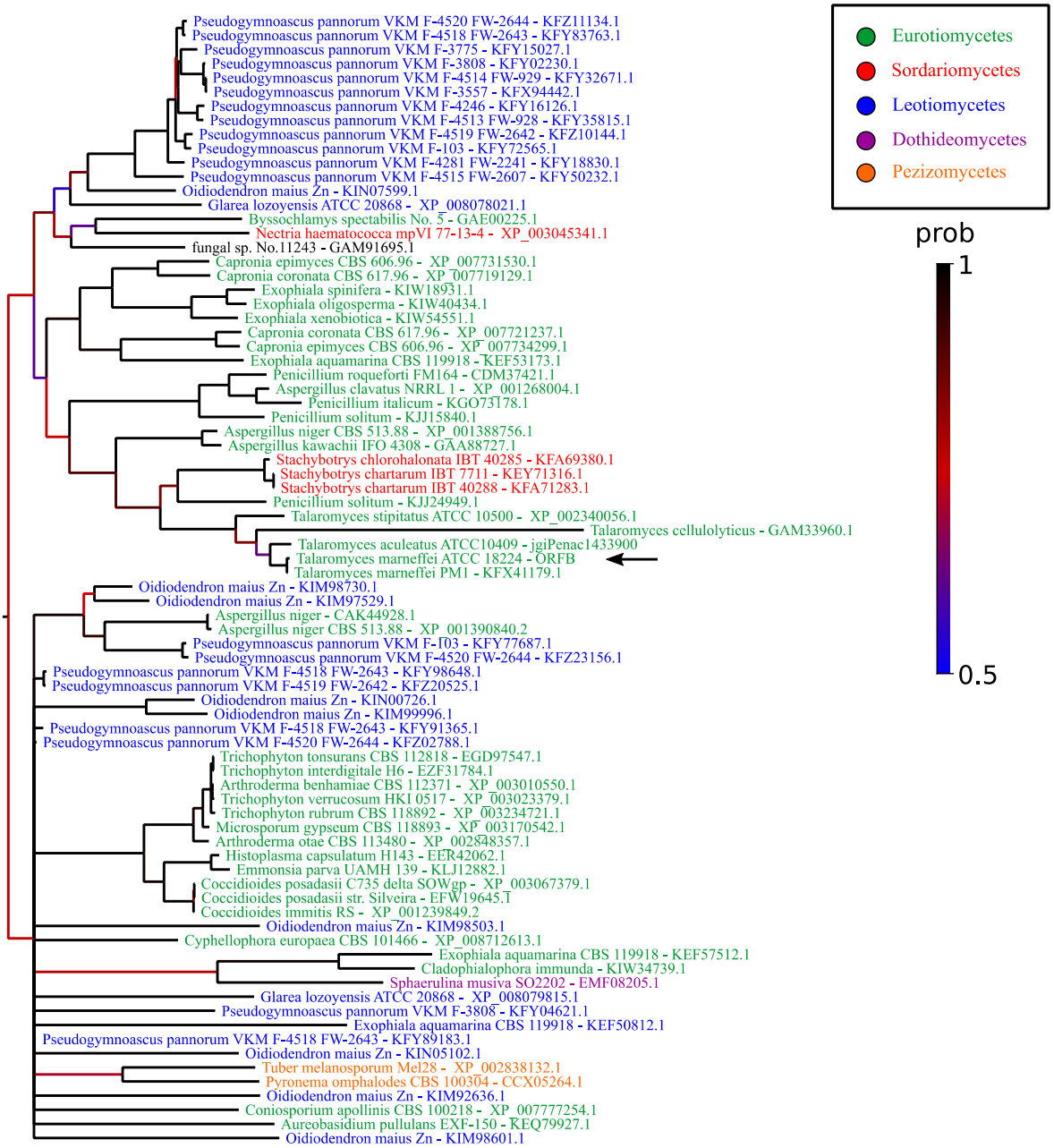


Figure E.8. Phylogenetic relationship of *T.* marneffei ATCC18224 ORFB homologs. The tree is a 50% majority rule consensus constructed with MrBayes. Branches are coloured according to its posterior probability, see colour gradient scale. Arrows point toward proteins coming from genes appearing in the azaphilone clusters in figure 7.1. Each protein has been coloured according to the taxonomical class of the species it is coming from (see figure legend).